

# Impact of Refractivity Profiles from a Proposed GNSS-RO Constellation on Tropical Cyclone Forecasts in a Global Modeling System

MICHAEL J. MUELLER,<sup>a</sup> ANDREW C. KREN,<sup>b</sup> LIDIA CUCURULL,<sup>c</sup> SEAN P. F. CASEY,<sup>b</sup>  
ROSS N. HOFFMAN,<sup>b</sup> ROBERT ATLAS,<sup>c</sup> AND TANYA R. PEEVEY<sup>d</sup>

<sup>a</sup> *University of Colorado Boulder, Cooperative Institute for Research in the Environmental Sciences, and NOAA/OAR/ESRL/Global Systems Laboratory, Boulder, Colorado*

<sup>b</sup> *Cooperative Institute for Marine and Atmospheric Studies, and NOAA/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida*

<sup>c</sup> *NOAA/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida*

<sup>d</sup> *Colorado State University, Cooperative Institute for Research in the Atmosphere, and NOAA/OAR/ESRL/Global Systems Laboratory, Boulder, Colorado*

(Manuscript received 5 November 2019, in final form 12 May 2020)

## ABSTRACT

A global observing system simulation experiment (OSSE) was used to assess the potential impact of a proposed Global Navigation Satellite System (GNSS) radio occultation (RO) constellation on tropical cyclone (TC) track, maximum 10-m wind speed ( $V_{\max}$ ), and integrated kinetic energy (IKE) forecasts. The OSSE system was based on the 7-km NASA nature run and simulated RO refractivity determined by the spatial distribution of observations from the original planned (i.e., including both equatorial and polar orbits) Constellation Observing System for Meteorology, Ionosphere, and Climate-2 (COSMIC-2). Data were assimilated using the NOAA operational weather analysis and forecasting system. Three experiments generated global TC track,  $V_{\max}$ , and IKE forecasts over 6 weeks of the North Atlantic hurricane season in the North Atlantic, east Pacific, and west Pacific basins. Confidence in our results was bolstered because track forecast errors were similar to those of official National Hurricane Center forecasts, and  $V_{\max}$  errors and IKE errors showed similar results. GNSS-RO assimilation did not significantly impact global track forecasts, but did slightly degrade  $V_{\max}$  and IKE forecasts in the first 30–60 h of lead time. Global forecast error statistics show adding or excluding explicit random errors to RO profiles made little difference to forecasts. There was large forecast-to-forecast variability in RO impact. For two cases studied in depth, track and  $V_{\max}$  improvements and degradations were traced backward through the previous 24 h of assimilation cycles. The largest  $V_{\max}$  degradation was traced to particularly good control analyses rather than poor analyses caused by GNSS-RO.

## 1. Introduction

As new observing systems have become available, their data have been incorporated into global weather models. In many cases, new global observing systems have improved model forecasts over the past several decades (Kalnay 2002). Despite these advances, further forecast improvements using new observing systems are not guaranteed. The decision to invest the resources required to build new observing systems should be well informed. For example, stakeholders should weigh the system's expected forecast impact against its expected

economic cost in comparison to similar cost benefit ratios of other proposed systems.

A good method to determine the forecast impact of proposed observations is through observing system simulation experiments (OSSEs). An OSSE is designed to mimic real-data denial studies [i.e., observing system experiments (OSEs)] with data from observing systems that do not yet exist (e.g., Atlas 1997; Hoffman and Atlas 2016). Data denial studies compare two experiments: A control generating forecasts from analyses made by assimilating all observations and a replica experiment excluding one observation type from assimilation. This makes OSSEs particularly suited to help stakeholders and forecasters anticipate forecast impacts of proposed observing systems.

Corresponding author: Michael J. Mueller, michael.mueller@noaa.gov

In 2017, the U.S. Congress mandated NOAA to use the OSSE framework to assess the impact of Global Navigation Satellite System radio occultation (GNSS-RO; e.g., [Businger et al. 1996](#); [Kursinski et al. 1997](#); [Hajj et al. 2002](#)) data assimilation on global weather forecasts ([Weather Research and Forecasting Innovation Act 2017](#), H.R. 353). Improvements in global forecasting were found during the experiments ([Cucurull et al. 2018](#)), supporting previous evidence of the positive impact of GNSS-RO observations on operational weather forecast skill (e.g., [Cucurull et al. 2007](#); [Anthes et al. 2008](#); [Buontempo et al. 2008](#); [Healy 2008](#); [Cucurull 2010](#); [Anlauf et al. 2011](#); [Cucurull and Anthes 2014](#); [Cucurull et al. 2017](#)).

Since tropical cyclones (TCs) are a leading cause of death and property damage, improving TC track and intensity forecasts is an area of active research. Previous studies exploring the impact of GNSS-RO data on TC analyses and forecasts have used the OSSE framework with regional models (e.g., [Huang et al. 2005](#); [Kuo et al. 2009](#); [Huang et al. 2010](#); [Liu et al. 2012](#); [Anisetty et al. 2014](#); [Chen et al. 2015](#); [Phunthirawurthi et al. 2016](#)). All but two of these papers looked at case studies of 1–2 TCs and their conclusions are therefore limited. The exceptions, [Huang et al. \(2010\)](#) and [Chen et al. \(2015\)](#), tested the impact of GNSS-RO data on track forecasts of several TCs.

[Huang et al. \(2010\)](#) used the Weather Research and Forecasting (WRF) limited-area model with nested 45–15–5-km grids to examine the impact of GNSS-RO observations from the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) satellites on TC track forecasts. They used different model domains for different TCs in the western Pacific and Indian Oceans, so the analysis and forecasting system was not standardized across cases. The study yielded mixed results with some cases showing large impacts with only a few GNSS-RO profiles and some cases with statistically insignificant impacts assimilating more profiles. The number of GNSS-RO profiles assimilated during the study was 2 to 56 per cycle over cases spanning 3 years and 13 forecasts. Their study of six typhoons in 2008 showed neutral impacts on track forecasts on the 45 km grid through 72-h lead time. Only 12 to 43 GNSS-RO profiles per cycle were assimilated during these six storms. [Chen et al. \(2015\)](#) used a limited area Typhoon WRF (TWRF) system at 45-km grid spacing to study GNSS-RO impact for 11 typhoons over the western Pacific Ocean from 2008 to 2010. Each assimilation cycle ingested about 100 GNSS-RO profiles. Results indicated that the GNSS-RO sounding assimilation reduced 72-h track errors by around 12 km, a modest

but statistically significant improvement. This improvement was credited to an improvement in the environmental steering current around the western Pacific subtropical high. Neither study evaluated any TC intensity metric (e.g., maximum wind speed;  $V_{\max}$ ). A comparison of these findings to the findings of our study can be found in [section 5](#).

There are three key reasons why GNSS-RO observations may be particularly beneficial for TC forecasts (e.g., [Bauer et al. 2014](#); [Ho et al. 2014](#); [Vergados et al. 2014](#); [Biondi et al. 2015](#); [Li et al. 2015](#); [Murphy et al. 2015](#)). First, since GNSS-RO observations are satellite based, profiles exist over open ocean where other vertical profiles are rare. Second, the GNSS-RO radio signals are minimally attenuated due to clouds and precipitation; thus, GNSS-RO observations can potentially be acquired very near to the surface in and below dense cloud cover. TCs often develop over oceanic data voids and are characterized by dense cloud cover, so GNSS-RO observations may provide unparalleled sampling of the TC environment. Even the TC itself, including its dynamically significant inner core could be observed with a sufficiently large constellation of RO receivers (e.g., [Leidner et al. 2017](#)). Third, since GNSS-RO observations are dependent on atmospheric density, they provide important information for understanding atmospheric thermodynamics. In practice, GNSS-RO retrievals or data assimilation infer high vertical ( $\sim 100$  m) resolution profiles of temperature, water vapor, and pressure. Nevertheless, the impact of GNSS-RO observations on TC forecasts may be limited by the coarse horizontal ( $\sim 300$  km) resolution of the observations relative to the finescale processes important during the life cycle of a TC. In addition, relatively coarse model resolution may limit the impact of GNSS-RO observations on analyses and forecasts of important TC structures if the model is incapable of resolving those structures.

To the authors' knowledge, this is the only study to evaluate the impact of GNSS-RO data on TC forecasts within a global OSSE. Therefore, the purpose of this study is to 1) evaluate the ability of such a global OSSE to reliably assess the impact of GNSS-RO observations on TC forecasts, 2) assess the impact of GNSS-RO assimilation on global model TC forecasts to the extent that the OSSE allows, and 3) provide suggestions for potential improvements in future similar studies.

This study is an extension of [Cucurull et al. \(2018\)](#), which is an examination of the impact of the Constellation Observing System for Meteorology, Ionosphere, and Climate-2 (COSMIC-2; [Cucurull et al. 2017](#)) observations on global and hemispheric weather pattern forecasts. The present study is unique in that it evaluates

RO impacts on TC forecasts from a global forecast model system cycling over a 6-week period.

In [section 2](#), the OSSE configuration and experimental design are described in detail. [Section 3](#) presents global track and  $V_{\max}$  error statistics along with an assessment of the reliability of the results in this OSSE. The  $V_{\max}$  statistics are compared to integrated kinetic energy statistics to provide an assessment of intensity and physical size and to evaluate the reliability of using  $V_{\max}$  as a key metric in this OSSE with relatively coarse resolution. [Section 4](#) provides an investigation of two analyses, representing an improvement and degradation due to RO assimilation. Finally, [section 5](#) provides a conclusion that evaluates our results in the context of previous studies and offers recommendations for future studies regarding implementation and interpretation of TC impacts in global model OSSEs.

## 2. Methodology

### a. OSSE background

There are several key components of OSSEs: 1) A *nature run* (NR), a free-running dynamical simulation of the atmosphere that stands in as the real atmosphere, 2) *Simulated observations* obtained from the NR using observation operators, and 3) an *experimental analysis and forecasting system*.

A reliable OSSE system should be validated in multiple ways. First, the NR climatology should be similar to that of the real atmosphere. Second, the differences between the modeling system that produces the NR and the analysis and forecasting system should approximate the differences between the real atmosphere and operational modeling systems. This requires that the NR be from a state-of-the-art numerical model and be of higher resolution than the numerical model used in the data denial experiments. Thus, while the experimental modeling system should be similar to an operational model, it should not be identical (e.g., the forecast model typically has a lower resolution than a true state-of-the-art operational model since the NR is “lower resolution” than the real atmosphere). Third, the system must be calibrated properly for the simulated observations. This means that observation errors, counts, and geographic and vertical distribution should be similar to operations and that their impact on analyses and forecasts be similar to their real-world counterparts. Further information on OSSEs can be found in [Hoffman and Atlas \(2016\)](#) and its supplemental material.

This study uses the Community Global OSSE Package (CGOP; [Boukabara et al. 2016](#)) in concert with NCEP’s Global Data Assimilation System/Global Forecast System

(GDAS/GFS) analysis and forecasting system. Precise details of our OSSE system follow.

### b. Nature run

The global OSSE experiments in this study use the National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO) Goddard Earth Observing System Model, version 5 (GEOS-5) global mesoscale nature run (G5NR; [Putman et al. 2015](#)). The G5NR is a ~2-yr dataset on a  $1/16^\circ$  (~7-km) horizontal grid and a 72-layer hybrid sigma-pressure vertical grid. The free-running nonhydrostatic GEOS-5 atmospheric general circulation model starts from real-world initial conditions in May 2005. Unconstrained to reality except through sea surface temperatures and sea ice from the daily  $0.05^\circ$  Operational Sea Surface Temperature and Sea Ice Analysis product ([Donlon et al. 2012](#)), the G5NR quickly develops a highly realistic state that is distinct from the real atmosphere. This study uses data 15 months into the free run, so there is no one-to-one relationship between individual real-world storms and those simulated in the G5NR.

The G5NR has been extensively documented and verified against the known climatology of the real atmosphere ([Gelaro et al. 2014](#)). While G5NR has been noted for realistic TC frequencies and tracks (e.g., [Gelaro et al. 2014](#); [Reale et al. 2017](#)), the algorithm used by [Reale et al. \(2017\)](#) to generate the G5NR TC climatology used strict warm-core and vertical-structure thresholds for determining the existence of a TC. Therefore, some weaker storms (e.g., tropical depressions) were missed, which skews the distribution of TC strength toward stronger storms. To somewhat alleviate this issue, we applied a TC genesis and lysis extension to each track from the original climatology; however, this did not apply to storms that were missed entirely by the [Reale et al. \(2017\)](#) climatology.

### c. Simulated observations

Simulated control observations for this study (surface pressure, temperature, zonal and meridional wind, specific humidity, and satellite radiances) were generated using appropriate forward operators following the procedures of [Boukabara et al. \(2016\)](#). Spatial and temporal coverage for conventional and radiance observations was based on real-world coverage from August to September 2014 ([Table 1](#)). Simulated conventional observations were generated by interpolating the nature run to the geographical locations of the corresponding real-world observations. Simulated clear sky satellite radiances were generated using the Community Radiative Transfer Model (CRTM), version 2.1.3. Geographic and temporal distribution of clear sky radiances were derived

TABLE 1. Observations assimilated during all tests conducted in this study. RO refractivity was not assimilated during CTL.

Observation type	Observation IDs	Description
Surface pressure ( $p_s$ )	120, 180, 181, 182, 187	Rawinsonde, surface marine, surface land, dropsonde, surface METAR
Wind ( $u, v$ )	220, 221, 223, 224, 229, 230, 231, 232, 233, 242, 243, 244, 245, 246, 250, 252, 253, 257, 258, 259, 280, 290	Rawinsonde, PIBAL, NPN wind profiler, NEXRAD, wind profiler-PIBAL decoded, aircraft, dropsonde, aircraft, JMA, EUMETSAT, NESDIS-GOES, MODIS-POES ( <i>Aqua</i> ), surface marine, ASCAT
Temperature ( $T$ )	120, 130, 131, 132, 133, 180, 182	Rawinsonde, aircraft, dropsonde, aircraft, surface marine, dropsonde
Moisture ( $q$ )	120, 132, 180, 182	Rawinsonde, dropsonde, surface marine
Radiance		<i>Metop-A</i> ( <i>HIRS4</i> , AMSU-A, MHS, IASI), <i>Metop-B</i> (AMSU-A, MHS, IASI), <i>GOES-15</i> (Sounders 1–4), <i>Suomi NPP</i> (ATMS, CrIS), <i>Aqua</i> (AIRS, AMSU-A), <i>N15</i> (AMSU-A), <i>N18</i> (AMSU-A, MHS), <i>N19</i> (AMSU-A, MHS), <i>F17</i> (SSMIS), <i>F18</i> (SSMIS), <i>M10</i> (SEVIRI)
RO refractivity	746–751, 752–757	COSMIC-2A, COSMIC-2B

from real-world observations and those locations that coincided with the G5NR clouds were rejected by the assimilation system's quality control checks. This impacted infrared radiances more than microwave radiances.

Simulated COSMIC-2 GNSS-RO refractivity profiles were produced using the forward operator described by Cucurull (2010). GNSS-RO observation coverage was calculated based on realistic orbits of the 12 originally planned COSMIC-2 satellites, with six receiving satellites (COSMIC-2B) in polar orbits and six (COSMIC-2A) in equatorial orbits. COSMIC-2A and COSMIC-2B were each expected to provide  $\sim 6000$  refractivity profiles per day. Figure 1 shows the assimilated RO coverage from COSMIC-2 over two 6-h time windows, with higher data density between  $45^\circ\text{N}$  and  $45^\circ\text{S}$  due to COSMIC-2A. This represents a large increase in profiles to  $\sim 3000$  profiles per 6-h assimilation cycle compared to the most comparable studies discussed in section 1 (Huang et al. 2010: 2 to 56 profiles; Chen et al. 2015:  $\sim 100$  profiles).

Since real-world observations have errors (e.g., instrumental and representativeness), a credible OSSE must ensure the error statistics associated with the simulated observations closely match those of a corresponding real-world OSE. When simulating observations from a nature run, there naturally will not be any instrument error since there are no real instruments. Thus, these are known as “perfect observations.”

To make observations more realistic, random Gaussian errors were added to all observations based on an iterative process described by Errico et al. (2013). Satellite radiances also have appropriate biases added as described by Boukabara et al. (2016). The process of adding random errors involves performing a “perfect observation” OSSE (i.e., one that assimilates observations with no errors added) and comparing observation increment (observation minus background and observation minus

analysis;  $O - B$  and  $O - A$ ) statistics for each observation type to those of a real-world OSE. Variances from those increments are used to update an estimate of added observation errors. Note that by comparing the perfect-observation OSSE to the OSE, we account for all error sources (e.g., instrument and representativeness) as one all-inclusive error. All observation errors in this study were simulated using this method. Note that the simulated

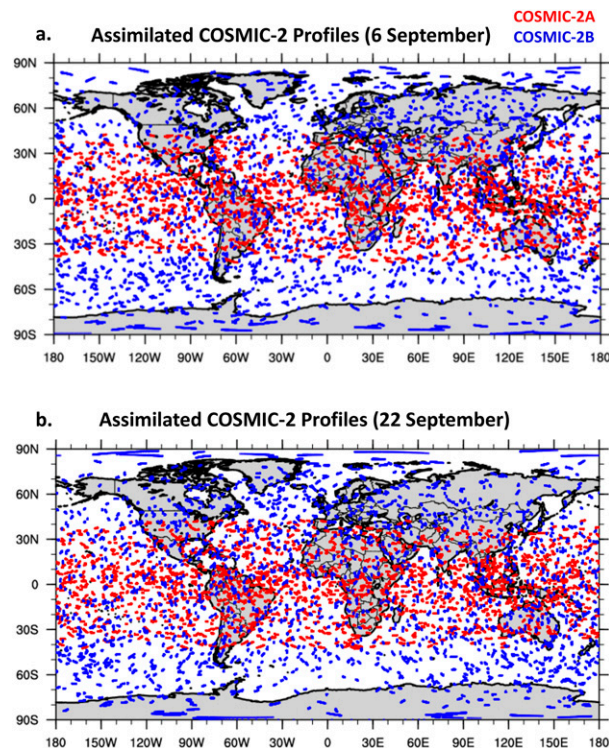


FIG. 1. Distribution of assimilated RO refractivity observations for the 0000 UTC cycles on (a) 6 Sep and (b) 22 Sep from COSMIC-2A (red) and COSMIC-2B (blue).



TABLE 2. Tests conducted to investigate RO impact on TC forecasts in the global OSSE. Descriptions include how each test is used in this study.

Test name	Observations assimilated	Description and use
CTL	$p_s, (u, v), T, q$ , radiance	Experiment cycled every 6 h for 2 months. Assimilates all operational observations except for RO soundings. Verification period: 15 Aug–30 Sep.
RO_err	$p_s, (u, v), T, q$ , radiance, RO refractivity (added-errors)	As in CTL, but assimilating RO refractivity every cycle. Compared with CTL, this is the primary test used to investigate accumulated RO impact on TC forecasts.
ROerr_1cyc_ctl	$p_s, (u, v), T, q$ , radiance	RO_err background + control observations. Compared with RO_err to investigate impact of current cycle's RO assimilation.
CTL_1cyc_ro	$p_s, (u, v), T, q$ , radiance, RO refractivity (added-errors)	CTL background + RO observations. Primarily used to ensure RO assimilating properly.
RO_perf	$p_s, (u, v), T, q$ , radiance, RO refractivity (perfect)	As in RO_err, but assimilating perfect RO observations. Compared with RO_err to explore the impact of added-errors on TC forecasts in a global model.

observation error statistics produced by this process are unique to the OSSE system being used. More detailed information on this method can be found in [Errico et al. \(2013\)](#), their appendix B).

This process was conducted for the 2014 operational configuration of the first COSMIC constellation. Increment variance statistics for COSMIC refractivity observations in the preliminary initial perfect observation OSSE and corresponding OSE are presented in [Cucurull et al. \(2018\)](#), Fig. 2a). The Gaussian random errors calculated for COSMIC are here added to COSMIC-2. This is justified because the instrument observation error differences between COSMIC and COSMIC-2 are expected to have very little if any impact. Also, measurement error is dwarfed by representativeness error in an OSSE configuration. The added errors are shown to be random and unbiased at all levels through an elevation of 30 km at four different latitudes in [Cucurull et al. \(2018\)](#), Fig. 3).

#### d. Analysis and forecast system

The global experiments used the Q1FY15 version of the National Centers for Environmental Prediction's Global Data Assimilation System (GDAS)/Global Forecast System (GFS) ([NWS 2014](#)). The analysis was created by a hybrid Gridpoint Statistical Interpolation (GSI)/ensemble Kalman filter (EnKF) configuration described by [Wang et al. \(2013\)](#) and [Kleist and Ide \(2015\)](#) (hybrid 3DVar). The GFS forecasts were run at reduced T670L64 resolution ( $\sim 27$  km at the equator). Eighty reduced-resolution T254L64 ( $\sim 77$  km at the equator) ensemble members were run to help estimate background error covariance via the EnKF ([Hamill et al. 2011](#); [Wang 2010](#); [Wang et al. 2013](#)). Final background error covariances used by GSI were generated by linearly combining the EnKF estimate and the static GSI estimate ([Boukabara et al. 2016](#)). Horizontal localization scales vary from 700 km in the troposphere to 1300 km toward the model top.

We conducted an 8-week (1 August–30 September of year 2 of the nature run simulation) cycling period with GDAS analyses produced four times per day at 0000, 0600, 1200, and 1800 UTC. The first two weeks of cycling (1–14 August) were used for model spinup and the remaining 6 weeks for verification purposes. GFS produced once-daily 168-h global forecasts initialized at 0000 UTC, yielding a total of 47 global forecasts. This once-daily (every 4 GDAS cycles) spacing helps mitigate forecast autocorrelation that arises due to the large influence of the previous analysis on the current one ([Aberson and DeMaria 1994](#)).

#### e. Experimental design

To examine the impact of GNSS-RO observations on TCs, we use the three global experiments conducted by [Cucurull et al. \(2018\)](#): 1) A control that assimilated all error-added conventional and satellite radiance observations (i.e., surface pressure, temperature, specific humidity,  $u$ - and  $v$ -wind, and satellite radiance) (CTL), 2) An experiment that assimilated all the observations used in CTL and perfect (i.e., error-free) GNSS-RO refractivity profiles (RO\_perf), and 3) An experiment that assimilated all the observations used in CTL and error-added GNSS-RO refractivity profiles (RO\_err). These global experiments tested the *accumulated* impact of 6-hourly GNSS-RO assimilation on TC track and  $V_{\max}$  forecasts. Supplemental “single-analysis” forecasts were run for select forecasts chosen for in-depth study (see [Table 2](#) for details on all tests).

#### f. Analysis and forecast verification

We assessed the average TC track and  $V_{\max}$  errors of all forecasts globally. Storm track and  $V_{\max}$  errors were computed using “best tracks” derived from 7-km G5NR data ([Figs. 2a–c](#)). (Errors are defined as the difference between the model state and the G5NR state.) The best

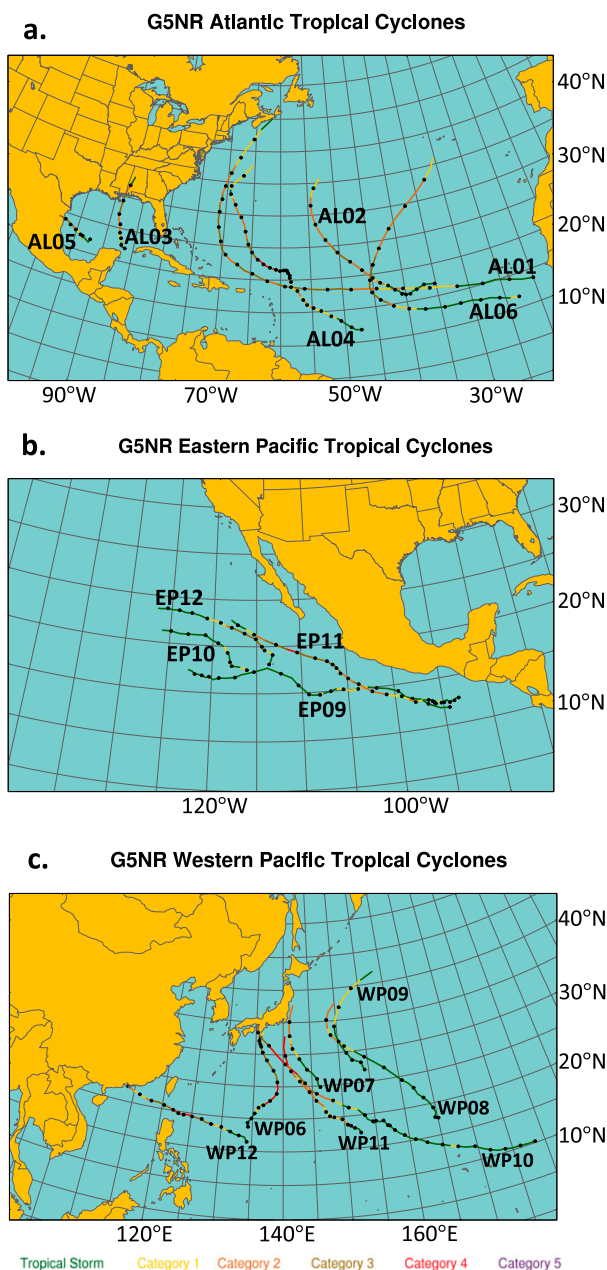


FIG. 2. G5NR tracks and categories for tropical cyclones in the (a) Atlantic, (b) eastern Pacific, and (c) western Pacific basins during the experimental period. Tracks based on Reale et al. (2017). Black dots represent location of cyclone center at 12 h intervals. Colors represent categories of  $V_{\max}$ .

tracks were based on the those published by Reale et al. (2017), extended to better capture genesis and lysis. Any G5NR TCs that began prior to or ended after the experimental period (15 August–30 September) were discarded as truncated events, yielding a total of 17 TCs for investigation. We used the Geophysical Fluid Dynamics Laboratory (GFDL) vortex tracker on  $0.25^\circ \times 0.25^\circ$  GFS

forecast data to provide track and  $V_{\max}$  forecasts for each forecast TC. Both G5NR and GFS tracks and minimum sea level pressure (MSLP) were cross checked against MSLP plots and both proved to be accurate. Both trackers have increments of  $0.1^\circ$  latitude/longitude for position and 1 kt ( $1 \text{ kt} \approx 0.51 \text{ m s}^{-1}$ ) for  $V_{\max}$ .

Since  $V_{\max}$  occurs in a small part of a TC and since the experimental analysis and forecasting system is coarse compared to the nature run, we also used a metric that assesses both maximum intensity and physical size. Integrated kinetic energy (IKE; Powell and Reinhold 2007) is defined as

$$\text{IKE} = \int \frac{1}{2} \rho U^2 dV, \quad (1)$$

where  $\rho$  is air density of  $1.15 \text{ kg m}^{-3}$ ,  $U$  is wind magnitude in excess of 34 kt,  $dV$  is the volume of the air under consideration ( $1 \text{ m}^3$ ), and IKE is expressed in terms of terajoules. IKE is the sum of kinetic energy of all unit volumes associated with the TC in which wind exceeds 34 kt. IKE is used to evaluate the impact of COSMIC-2 radio occultation and to provide a baseline of comparison for the  $V_{\max}$  metric. To provide a direct comparison both G5NR and experiment IKE values were computed on the GFS native  $0.25^\circ$  grid.

To quantify the error differences of storm track and intensity, a paired Student's  $t$  test is performed, which seeks to determine whether the observed differences in the experiments, on average, are significantly different (Hamill 1999; Gilleland et al. 2018). (Sample sizes for these  $t$  tests are given in the middle of Figs. 4, 5, and 6.) From this, the statistical significance of the mean differences between the experiments is determined using a 95% confidence interval, as in Kren et al. (2018).

### 3. Impact of COSMIC-2 on global track and intensity forecasts

#### a. Global impact on TC analyses

Globally, there were 17 TCs identified in the Reale et al. (2017) climatology dataset during the experimental period investigated in this study. There were 132 individual storm initializations within the 47 168-h GFS forecasts made during the study period. (This is possible because many of the global forecasts contained multiple TCs). Figure 3 shows histograms of the difference between CTL and RO\_err track and  $V_{\max}$  error magnitudes at the initial time of these storm forecasts. The histograms show that RO\_err track analyses skew slightly toward degradations, with a mean 4 km degradation (Fig. 3a). This falls well within the increment of the TC tracker ( $0.1^\circ$  or 9–12 km).

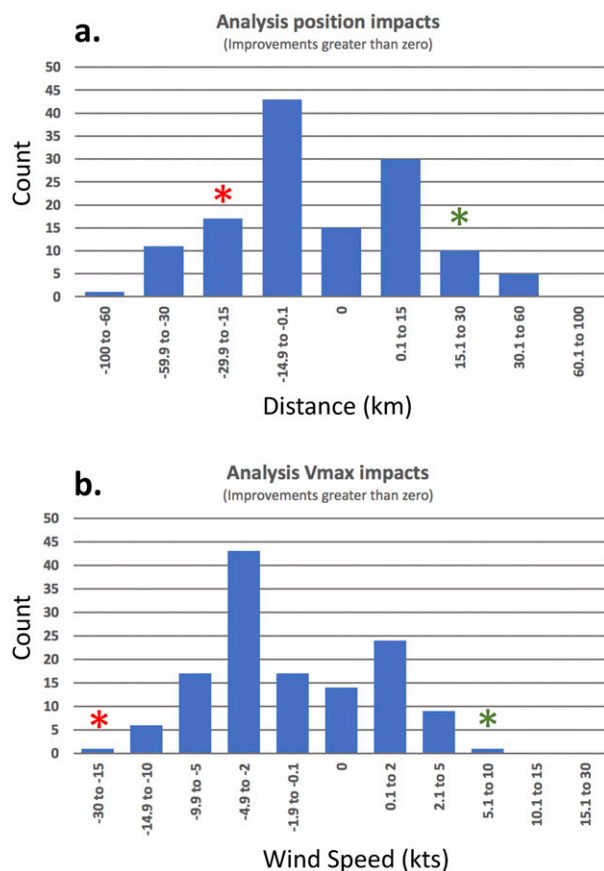


FIG. 3. Histograms of GFS 0h impact distribution for (a) track and (b)  $V_{\max}$ . Impact is defined as CTL error magnitude minus RO\_err error magnitude. Positive values denote RO\_err is improved compared to CTL. Red asterisk represents the degraded analysis at 0000 UTC 22 Sep for AL04; green asterisk represents the improved analysis at 0000 UTC 6 Sep for EP11.

RO\_err  $V_{\max}$  analyses skew more convincingly toward degradations, with a mean degradation of 2 kt (Fig. 3b). The TC tracker has  $V_{\max}$  resolution of 1 kt. For context, official TC Vitals information such as that from the HURDAT2 Atlantic hurricane database (Landsea and Franklin 2013) is published with an increment of 5 kt. Out of 132 analyses, only 25 (~19%) had impacts of 5 kt or larger, but those that did were overwhelmingly degraded (24 analyses) rather than improved (1 analysis).

#### b. Global impact on TC forecasts

Global analyses yield CTL track forecast errors that are very similar to those of the National Hurricane Center (NHC) for storms in the Atlantic basin during the 2014 hurricane season (Figs. 4a,c; NHC statistics can be found at NOAA 2019a,b). This enhances our confidence that the OSSE system used here adequately represents the impact of the operationally assimilated

observations on TC track forecasts. Both CTL and RO\_err track errors grow over time at a realistic rate through 120 h (the end of the NHC statistics).

The impact of COSMIC-2 RO refractivity profiles on track forecasts is less than 15 km at most lead times. Although the impacts are degradations at most lead times, only lead times 0, 6, and 24 h are statistically significant (4–10 km degradations). The magnitude of these degradations is smaller than the smallest possible difference between two individual TC positions due to tracker position increments ( $0.1^\circ$  or 9–12 km). Furthermore, the tracker is applied on a GFS forecast grid of  $0.25^\circ$  (~27 km). Thus, RO assimilation yields neutral track forecast impacts in this global model OSSE.

Global analyses yield CTL  $V_{\max}$  forecast errors that are 3–5 $\times$  larger those of the NHC in the Atlantic basin during the 2014 hurricane season (Figs. 4b,d). There are a few likely reasons for this: 1) The G5NR TC tracker was strict and therefore stronger TCs are over-represented in the sample; 2) the  $V_{\max}$  in our experiments is precisely known from the nature run, while it can only be approximated in real storms; 3) the experimental analysis and forecast modeling system has a lower-than-operational resolution (~77 and ~27 km at the equator, respectively); and 4) NHC forecasts are not model products and are based in part on observations that sample scales of motion not represented in the GDAS, GFS, or G5NR. Ultimately, this reduces confidence that the  $V_{\max}$  metric presented here can be used as a direct comparison to future real-world data impact experiments; however, there is value in documenting these impacts in the global OSSE system. Specifically, it is valuable to provide future studies with a point of comparison for their own results, and to act as an indicator of structural changes in the storm. To further enhance our analysis, we also calculated integrated kinetic energy (IKE) as an intensity and physical size metric and compared IKE and  $V_{\max}$  results.

Average  $V_{\max}$  errors of 35–45 kt characterized both CTL and RO\_err. The degradations seen in RO\_err analyses extended through 60 h forecast lead time, characterized by 1–2 kt statistically significant degradations. Note in Fig. 4d that the degradation at 0 h does not grow; instead it remains ~2 kt until 48 h, at which point it begins to decrease. This strongly suggests the source of the errors in the first 60 h is not in the broader environment. Rather, the source of the  $V_{\max}$  degradation seems to be in the analysis of the storm itself. Section 4 explores this in more detail.

Global IKE errors are similar to those of  $V_{\max}$  (Figs. 5a,b). Statistically significant degradations of around 1 TJ occur during the first 30 h of lead time. Like  $V_{\max}$ , the degradation does not grow and impact becomes neutral

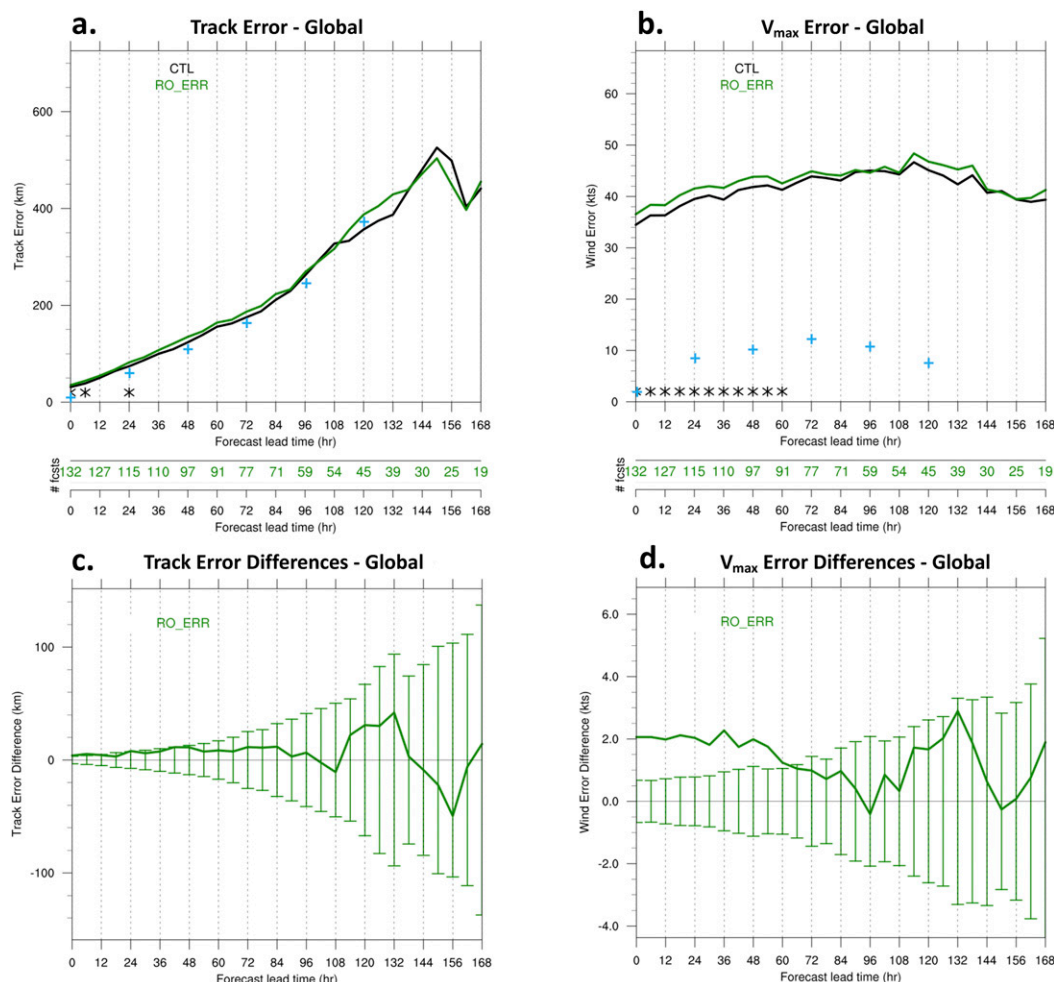


FIG. 4. Global average (a) track and (b)  $V_{\max}$  errors for CTL and RO\_err. Black (green) asterisks along x axis denote statistically significant degradations (improvements). RO\_err (c) track and (d)  $V_{\max}$  differences with respect to CTL. Vertical bars in (c) and (d) are confidence intervals at the 95% level. Blue plus signs denote average National Hurricane Center forecast errors for the Atlantic basin in 2014.

after 30h. Compared to  $V_{\max}$ , significant degradations last only half the time (30 versus 60 h). Unlike  $V_{\max}$ , the IKE error for both tests grow throughout the 168 h forecasts. Since this metric that accounts for physical size of the TC supports the results found with  $V_{\max}$ , RO assimilation yields small degradations in the first 30–60 h of forecasts in this global model OSSE. Furthermore, confidence increased that the conclusions drawn from  $V_{\max}$  results are appropriate for assessing impact in this global model OSSE.

Adding simulated observation errors does not alter RO impacts in a statistically significant way, though RO\_perf tends to produce smaller errors than RO\_err (Fig. 6). It was expected and encouraging that RO\_perf produces smaller forecast errors than RO\_err. The lack of statistically significant differences suggests adding random errors is a second-order impact and is not vital to producing

realistic TC forecast errors in a global model OSSE. The obvious exception to this is significant degradation due to adding observational errors in  $V_{\max}$  forecasts at 78 and 114 h, which slightly exceeded the 95% confidence interval.

As in the Huang et al. (2010) experiments discussed in the Introduction, there was strong forecast-dependency (i.e., some of the 132 forecasts were improved and others degraded). An attempt was made to identify storm attributes that distinguish between forecasts that benefited from RO assimilation from those that did not. We binned TCs at initialization by basin, latitude band,  $V_{\max}$ , minimum sea level pressure, and RO observation coverage in the vicinity of the TC core (not shown). This binned approach of course yielded smaller samples, and therefore often yielded few statistically significant differences between CTL and RO\_err. Where significant differences were found, small forecast sample size hindered



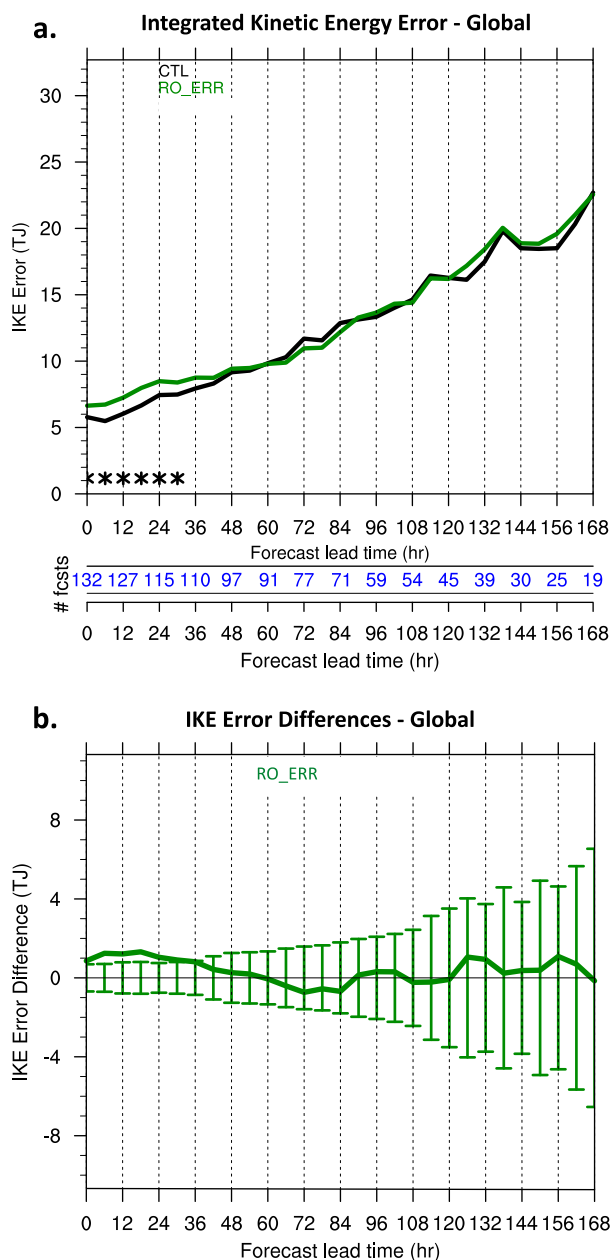


FIG. 5. As in Fig. 4, but for integrated kinetic energy (IKE).

the ability to determine conclusively whether those differences were caused by the attribute under study or whether it was the result of a deficiency of the sample. No overall distinguishing attribute was found through binning the results of this study.

#### 4. GNSS-RO impact on two forecasts

In light of these global results and the degradation of the  $V_{\max}$  and IKE forecasts, in-depth study of two individual forecasts was conducted. The forecasts were

chosen based on the *accumulated* impact of RO assimilation on analyses. Accumulated impact is due to the combined effect of 1) the current-cycle impacts and 2) the impact of RO on the background field acquired over all previous cycles. This is a confounding factor for any realistic cycling data impact study, but especially one that investigates impact on individual weather systems (such as TCs) which can be more sensitive to initial conditions than time-averaged global metrics.

To partially address this complication and determine which impacts are due to current-cycle observations and which are due to the accumulation of impacts in the background field, we ran single-cycle tests assimilating control observations and an RO\_err background (ROerr\_1cyc\_ctl). This forecast is compared to CTL and RO\_err to determine the contribution of 1) accumulated impact in the background field and 2) single cycle impacts due to RO profiles withheld in the current 6-h window. Note that ROerr\_1cyc\_ctl provides an imperfect picture of RO impact in a given assimilation window because not assimilating RO changes how the other observations are assimilated. Therefore, the impact seen in these comparisons may not be the same as the impact seen through a more direct comparison of RO\_err background and analyses. This second approach was also employed in this study and is presented later.

The assimilation itself is evaluated using statistics generated by single-cycle tests assimilating control and COSMIC-2 observations and a CTL background (CTL\_1cyc\_ro). Since cycling RO\_err is more representative of how RO would be used in operations, CTL\_1cyc\_ro is not meant to show realistic RO impact.

##### a. Forecast and assimilation overview

To produce a clear contrast between forecasts, we chose the most extreme  $V_{\max}$  analysis impacts (see Fig. 3): 6 September forecast of EP11 (improvement) and 22 September forecast of AL04 (degradation). These cases also had track analysis impacts of the same sign. IKE calculations for these two analyses and forecasts were similar to  $V_{\max}$  results described in this section (not shown).

The track forecasts for EP11 show CTL, RO\_err, and ROerr\_1cyc\_ctl move to the right of the G5NR track through 96 h, resulting in landfall along the southwestern coast of Mexico (Fig. 7a). In RO\_err, there was modest improvement in TC center initialization, followed by steady improvement of 30–50 km after 12 h (Fig. 7c). ROerr\_1cyc\_ctl was similar to RO\_err as expected, showing that accumulation of impacts in the background state was considerably larger than the current cycle. Due to model resolution limitations,  $V_{\max}$  forecasts for each test fell 20–25 kt short of G5NR

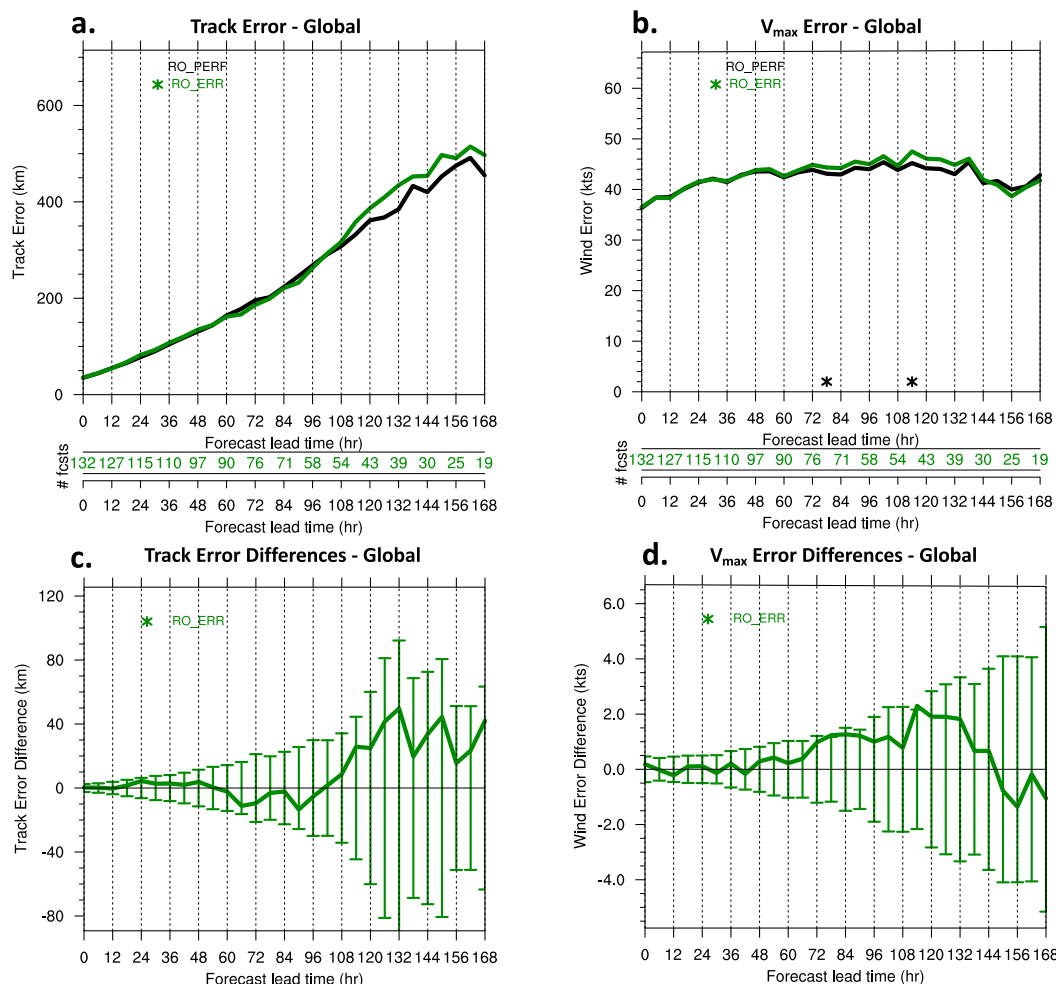


FIG. 6. As in Fig. 4, but comparing RO\_err to RO\_perf.

(Fig. 7b). RO\_err and ROerr\_1cyc\_ctl performed best at  $V_{\max}$  initialization and for much of the first 72 h. Analysis improvements did not increase during the forecast.

The CTL\_1cyc\_ro analysis was very similar to CTL (expected because CTL\_1cyc\_ro used CTL background), but consistently produced much better track forecasts than the other three tests through 84 h. CTL\_1cyc\_ro  $V_{\max}$  impact through 48 h was small, but it produced the best (strongest  $V_{\max}$ ) forecast thereafter likely due to less land interaction allowed by its superior track forecast. Overall, the assimilation of RO into the CTL background did not appear to degrade the forecast in a way that would suggest poor assimilation.

The track forecasts for AL04 show CTL, RO\_err, and ROerr\_1cyc\_ctl move to the right of G5NR through 72 h (Fig. 8a). RO\_err and ROerr\_1cyc\_ctl initialize the TC position poorly (see Fig. 12) and that degradation grows steadily over the first 48 h. The degradation in RO\_err and ROerr\_1cyc\_ctl was 100–250 km from 12 to 96 h.

After 96 h, the G5NR storm curves sharply to the northeast. ROerr\_1cyc\_ctl did not meet TC criteria at 120 h and was therefore terminated in the tracker. RO\_err track error grows rapidly to 1700 km by 132 h, a result of failure to adequately curve. CTL forecasted the curve better, producing 1400 km improvements at 132 h. Again,  $V_{\max}$  initialized poorly in the coarse global model (50–70 kt underestimate) in all three tests (Fig. 8b). RO\_err and ROerr\_1cyc\_ctl each degraded  $V_{\max}$  by 18 kt, suggesting the accumulated impact of RO assimilation in the analysis and modeling system was much larger than the immediate impact of the current cycle of RO assimilation. Forecast error declined for all three tests as the G5NR TC weakened.

CTL\_1cyc\_ro analysis and forecast errors again were quite similar to CTL, as expected. The main difference was that the CTL\_1cyc\_ro track curved more sharply, leading to 100–200 km improvements over CTL after 114 h. As in the 6 September EP11 analysis and forecast, there was no

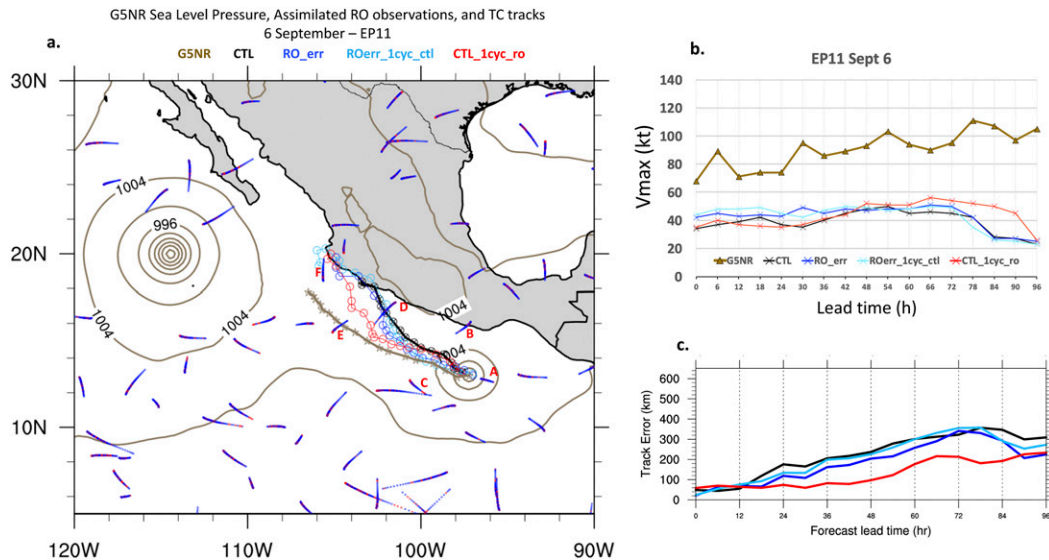


FIG. 7. Forecasts from CTL, RO\_err, ROerr\_1cyc\_ctl, and CTL\_1cyc\_ro for 6 Sep in the EP11 region. (a) GFS forecast tracks (markers every 6 h) with smoothed G5NR sea level pressure (valid at 0000 UTC 6 Sep) and RO refractivity observations assimilated in CTL\_1cyc\_ro. Red observations denote those with larger  $O - A$  bias than  $O - B$  bias; blue denotes smaller  $O - A$  bias than  $O - B$  bias. Individual profiles denoted A–F featured in Fig. 9. (b)  $V_{max}$  plots for G5NR, CTL, RO\_err, ROerr\_1cyc\_ctl, and CTL\_1cyc\_ro. (c) Track error magnitude by lead time.

indication assimilating RO into the CTL background caused degradations indicative of poor assimilation.

Before evaluating how RO assimilation caused these impacts, it was vital to further ensure GSI assimilated RO well at these cycles. Assimilated RO observations that are plotted in blue represent  $O - A$  bias that was smaller than  $O - B$  bias (Figs. 7a and 8a) in CTL\_1cyc\_ro. Observations plotted in red represent  $O - A$  bias that

was larger than  $O - B$  bias in CTL\_1cyc\_ro. Over 70% of RO observations in CTL\_1cyc\_ro had a better fit to the model after assimilation both globally and in the regions defined in Figs. 7a and 8a. Assimilated RO counts, bias, and standard deviation for both CTL\_1cyc\_ro and RO\_err are shown in Table 3. This shows reductions in bias and standard deviation during assimilation, a sign that GSI was using RO observations to successfully

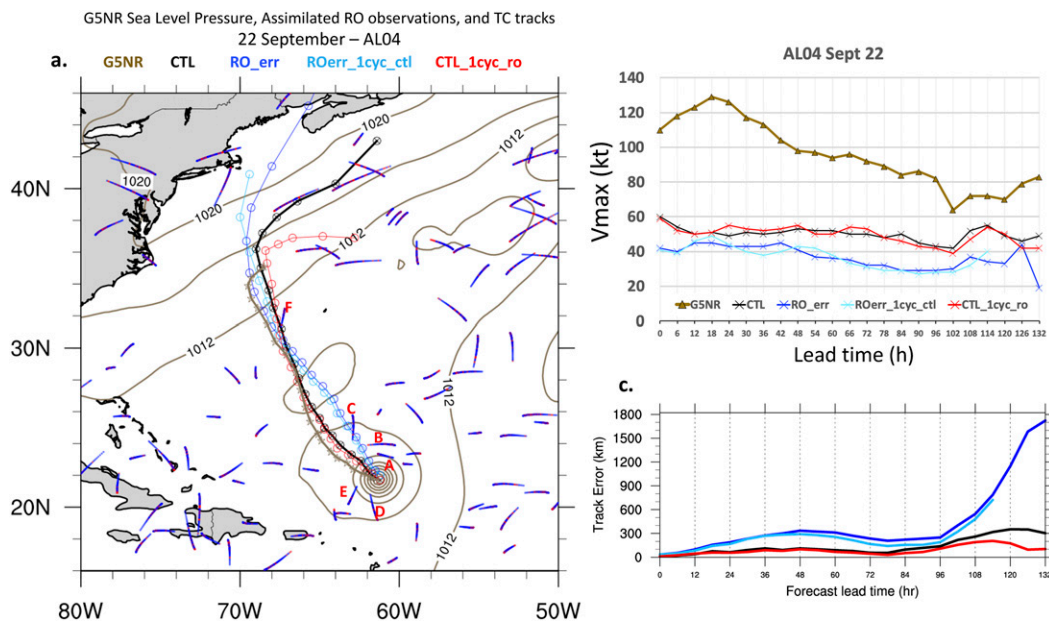


FIG. 8. As in Fig. 7, but for 22 Sep in the AL04 region. Individual profiles denoted A–F featured in Fig. 10.

TABLE 3. Assimilation counts, bias, and standard deviation for RO\_err and CTL\_1cyc\_ro on global and regional domains. Bias and standard deviation are of incremental refractivity  $100 \times [(O - G)/O]$ , a percentage difference between the observed and estimated (analysis or background) refractivity. AL04 and EP11 domains are the same as those shown in Figs. 7a and 8a.

	Assimilation count $O - B$ ; $O - A$	Bias $O - B$ ; $O - A$	Standard deviation $O - B$ ; $O - A$
6 Sep–global			
RO_err	517 601; 526 705	−0.006; −0.006	0.822; 0.459
CTL_1cyc_ro	511 345; 525 472	−0.015; 0.001	0.952; 0.480
22 Sep–global			
RO_err	511 681; 520 901	−0.010; −0.005	0.837; 0.466
CTL_1cyc_ro	506 295; 519 789	−0.027; 0.002	0.953; 0.480
6 Sep–EP11 region			
RO_err	8913; 9139	−0.077; −0.019	1.104; 0.489
CTL_1cyc_ro	8733; 9078	−0.154; −0.018	1.225; 0.515
22 Sep–AL04 region			
RO_err	12 107; 12 483	0.012; −0.008	1.053; 0.460
CTL_1cyc_ro	11 986; 12 453	−0.016; 0.004	1.135; 0.478

minimize  $O - A$  during these cycles (statistics for all cycles in RO\_err were similar after the 1–14 August spinup period). As expected, RO\_err O-B bias and standard deviation were smaller than CTL\_1cyc\_ro because RO\_err background included RO impact from previous cycles.

Further analysis shows background errors for individual profiles (denoted in Figs. 7a and 8a) in the vicinity of the TCs under study (Figs. 9 and 10). Profile A for each case is the profile that samples closest to the TC core. All profiles correct misfits between observation and background especially in the mid- to upper troposphere ( $\sim 2$ – $10$  km) and show no indication of abnormal assimilation near the TCs.

#### b. Impact on TC track and $V_{\max}$ analyses

A closer view of sea level pressure and 10-m wind speed analyses shows the impact of accumulated and single-cycle RO assimilation (Figs. 11 and 12). GFS TCs are weaker than G5NR storms and are generally not as symmetrical. While GFS TC cores are much weaker, the GFS storms are just as broad. For instance, the AL04 tropical storm force wind field is comparable in size for G5NR and CTL, but the core of the storm is much less compact in the GFS tests. Thus, these GFS storms do not feature hurricane force winds. RO\_err AL04 is displaced northward from G5NR and produces a much less symmetrical storm than CTL (degradation), while RO\_err EP11 is displaced eastward toward G5NR and has a stronger wind field than CTL (improvement). ROerr\_1cyc\_ctl is slightly weaker than RO\_err for AL04 (i.e., the TC was slightly improved due to RO assimilation at that cycle) and slightly stronger for EP11 (i.e., the TC was slightly degraded due to RO assimilation at that cycle). Note that the single cycle impacts

seen by comparing RO\_err and ROerr\_1cyc\_ctl are not the same as comparing RO\_err background and RO\_err analysis because the difference between assimilating control observations plus RO (as in RO\_err) and assimilating only control observations (as in ROerr\_1cyc\_ctl) is not only in the RO observations. The “control observations” in RO\_err and ROerr\_1cyc\_ctl are slightly different because adding RO impacts how the control observations are assimilated. The more direct analysis of comparing RO\_err background to RO\_err analysis is presented later.

The origins of the accumulated impacts can be traced back through previous cycles of analyses and short-term forecasts. Note again that this is different from the RO\_err/ROerr\_1cyc\_ctl comparison because it is comparing the impact of RO observations on the RO\_err background to the impact of control observations on the CTL background. Hence, this allows us to follow the evolution of the differences between CTL and RO\_err analyses leading up to the 0000 UTC forecasts of AL04 and EP11. Time series of analysis impacts for both storms reveal little correlation of impact sign or magnitude (Fig. 13). For example, Figs. 13c and 13d show abrupt changes in sign and/or magnitude in the cycles immediately preceding the outlier  $V_{\max}$  analyses.

Assimilation cycles occurred every 6 h, presenting an opportunity to show how the RO\_err analyses and short forecasts evolved over the previous day in comparison to that of CTL. Track forecast improvement for EP11 can be attributed to the background state at 0000 UTC 6 September since track error was not reduced in RO\_err through data assimilation at that cycle (Fig. 14a). This improvement can be traced back to good analyses at 1200 and 1800 UTC 5 September. While CTL analyses were also improved through assimilation, the 6-h GDAS



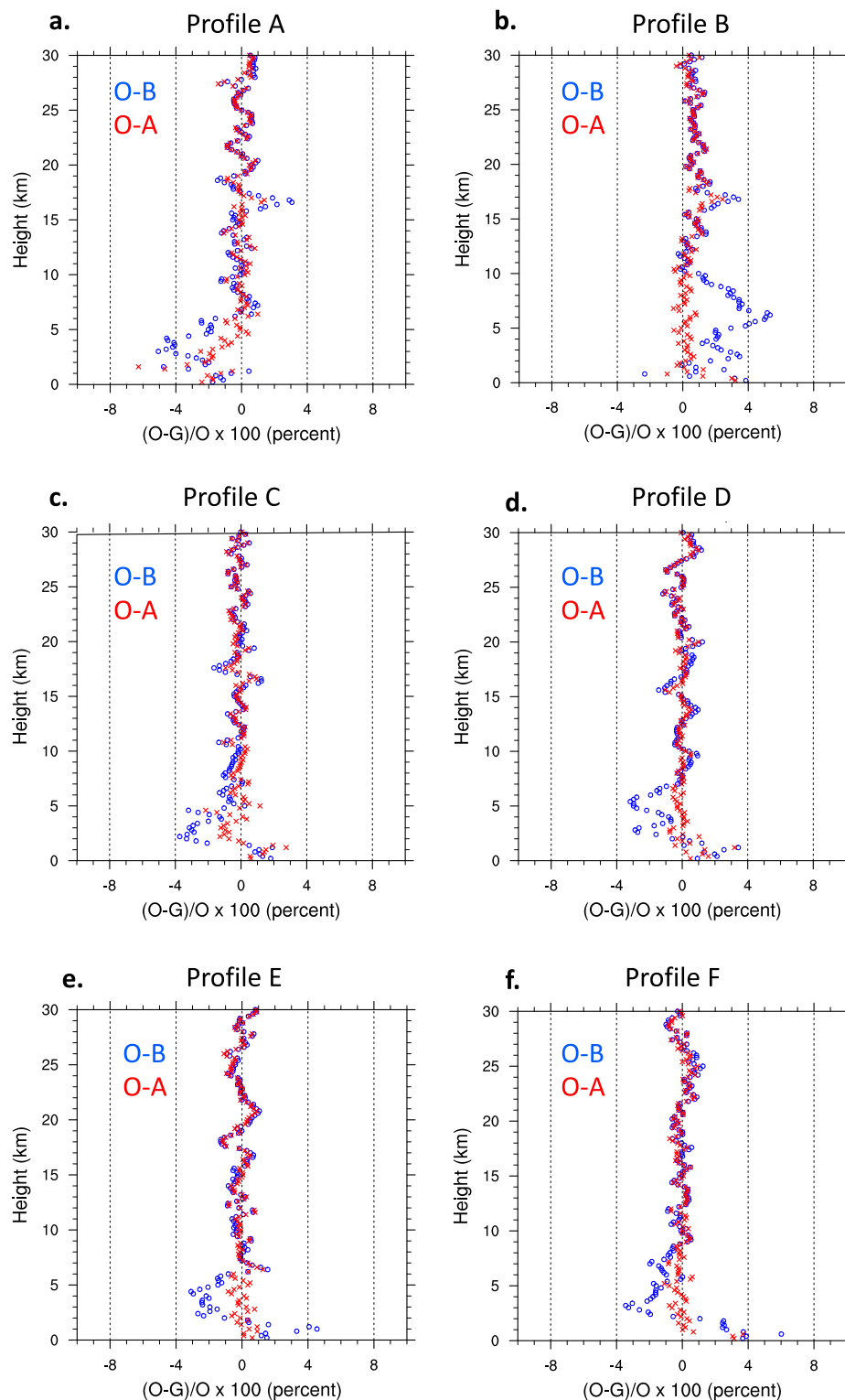


FIG. 9. Vertical profiles of  $O - B$  (blue) and  $O - A$  (red) bias for Profiles A–F on 6 Sep denoted in Fig. 7. Analysis and background are from the CTL\_1cyc\_ro test.

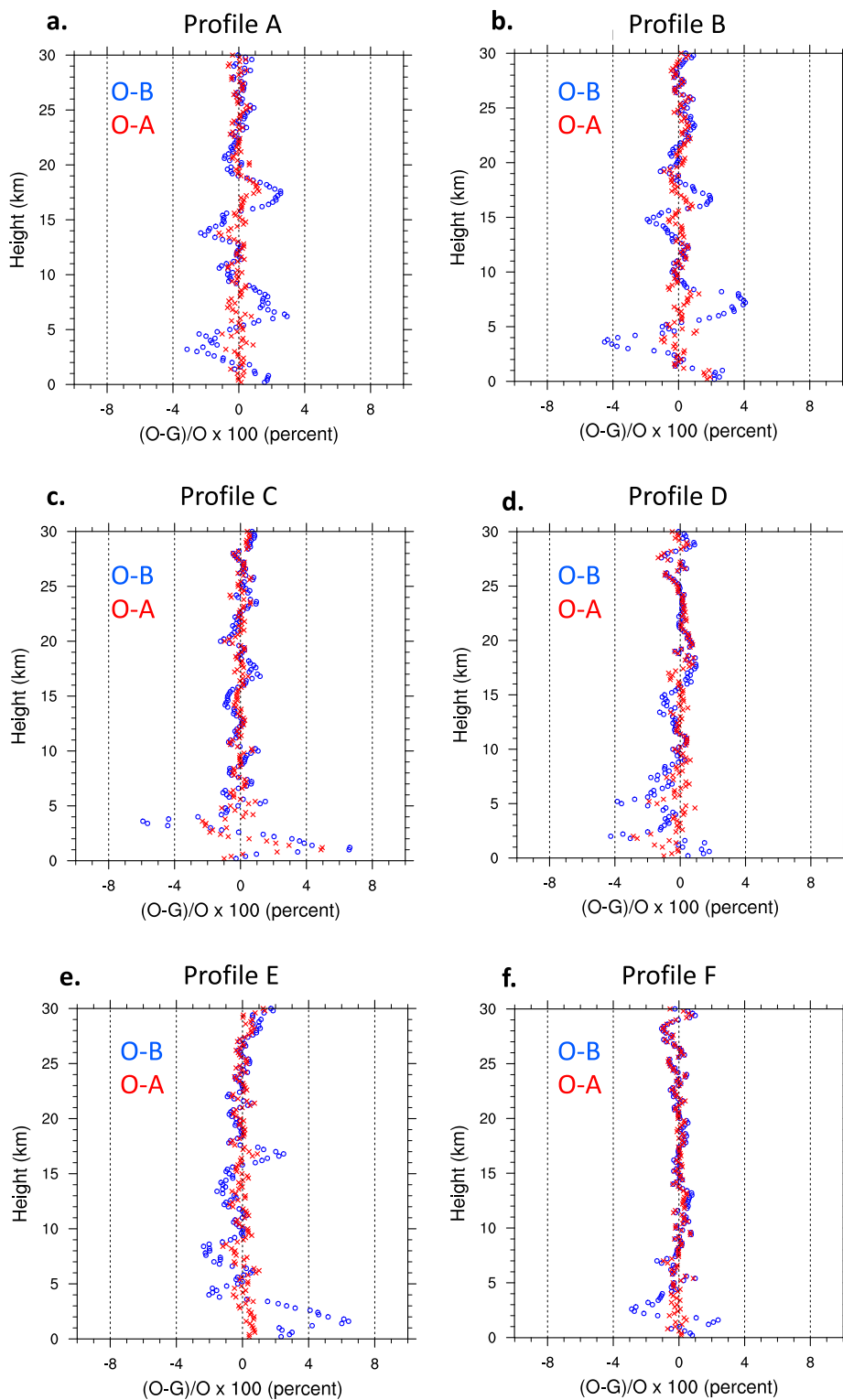


FIG. 10. As in Fig. 9, but for Profiles A–F on 22 Sep denoted in Fig. 8.

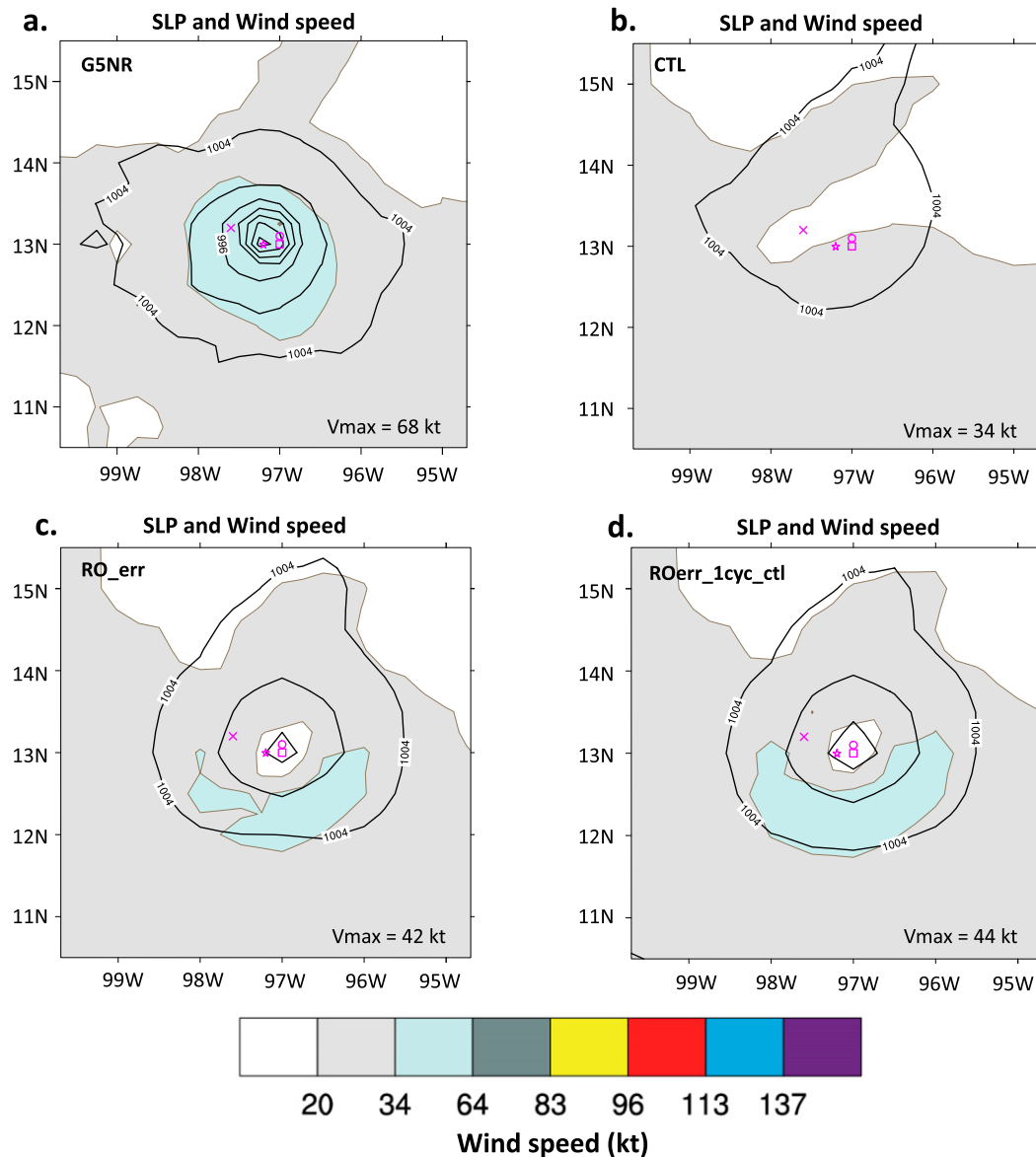


FIG. 11. Sea level pressure (contours) and 10-m wind speed (fill) for EP11 at 0000 UTC 6 Sep for (a) G5NR, (b) CTL, (c) RO\_err, and (d) ROerr\_1cyc\_ctl. Magenta markers denote TC tracker positions: G5NR: star, CTL: "x," RO\_err: square, ROerr\_1cyc\_ctl: "O."

forecasts yielded very poor background states. Thus, the 27 km track improvement for RO\_err was largely caused by a series of comparatively good 6-h forecasts. The track forecast degradation for AL04 can be attributed to a series of very poor 6-h forecasts at 0600, 1200, and 1800 UTC 21 September (Fig. 14b). Although the assimilation at 1200 and 1800 UTC 21 September and 0000 UTC 22 September greatly reduced the error, it was not enough to overcome the aforementioned poor forecasts. Meanwhile, the 6-h forecast initialized at 1200 UTC in CTL reduced the error from 46 to 25 km. Thus, the 25 km degradation was caused by poor 6-h

forecasts in RO\_err and one particularly good 6-h forecast in CTL. Note that while the improved EP11 analysis did not feature an improvement in the 0000 UTC 6 September assimilation cycle, the degraded AL04 analysis did feature an improvement in the 0000 UTC 22 September assimilation cycle. It is important to remember in a cycling experiment that a degradation at a particular cycle does not imply that cycle's assimilation caused the degradation.

For the two forecasts investigated here, G5NR  $V_{\max}$  was always higher than the GFS experiments, so any increase in GDAS winds constituted an improvement.

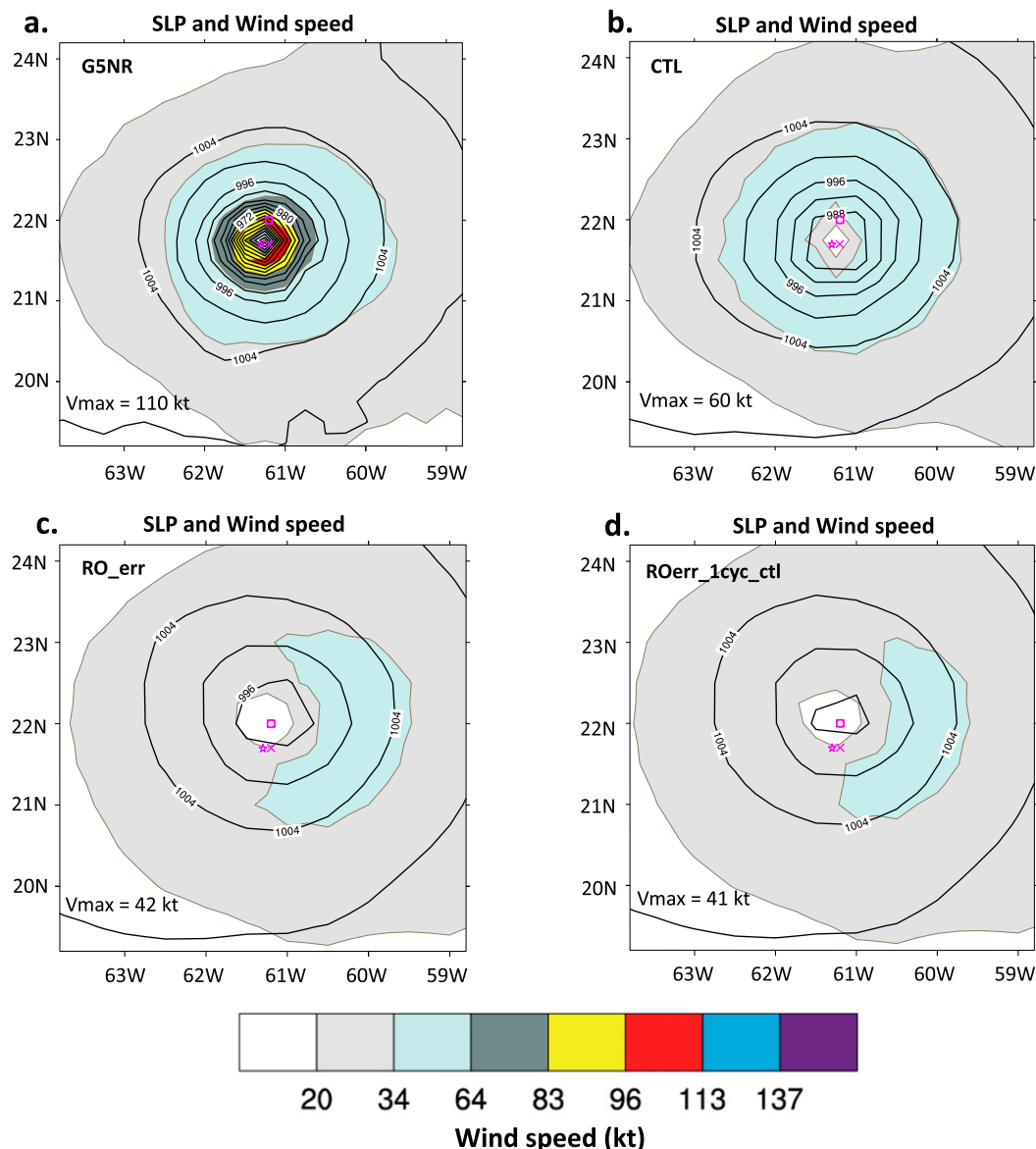


FIG. 12. As in Fig. 11, but for AL04 at 0000 UTC 22 Sep. Note that RO\_err and ROerr\_1cyc\_ctl are collocated.

The 8 kt improvement of EP11  $V_{\max}$  analysis at 0000 UTC 6 September developed during the 6-h forecast from 1800 UTC 5 September to 0000 UTC 6 September (Fig. 15a). The assimilation at the 1800 UTC cycle actually degraded  $V_{\max}$  by 3 kt. During the subsequent 6-h GDAS forecasts, RO\_err  $V_{\max}$  increased 7 kt and CTL  $V_{\max}$  decreased 3 kt. Together with a 3 kt weakening in the CTL assimilation and 1 kt strengthening in the RO\_err assimilation, this led to the observed 8 kt improvement. The 18 kt degradation in the AL04 analysis originated at the 1200 UTC cycle on 21 September (Fig. 15b). Two lackluster RO\_err analysis cycles followed from that time with  $V_{\max}$  unable to increase above

42 kt. Meanwhile, CTL analyses and 6-h GDAS forecasts generated progressively increasing  $V_{\max}$  to 60 kt, accounting for the 18 kt degradation in RO\_err.

## 5. Conclusions

While several studies have explored the impact of real RO profiles on TC forecasts in limited area models (e.g., Huang et al. 2010; Chen et al. 2015), to the authors' knowledge a similar investigation of simulated refractivity profiles in a global modeling system has not been published. Hence, this study was designed to 1) evaluate the ability of a global model OSSE to reliably assess the



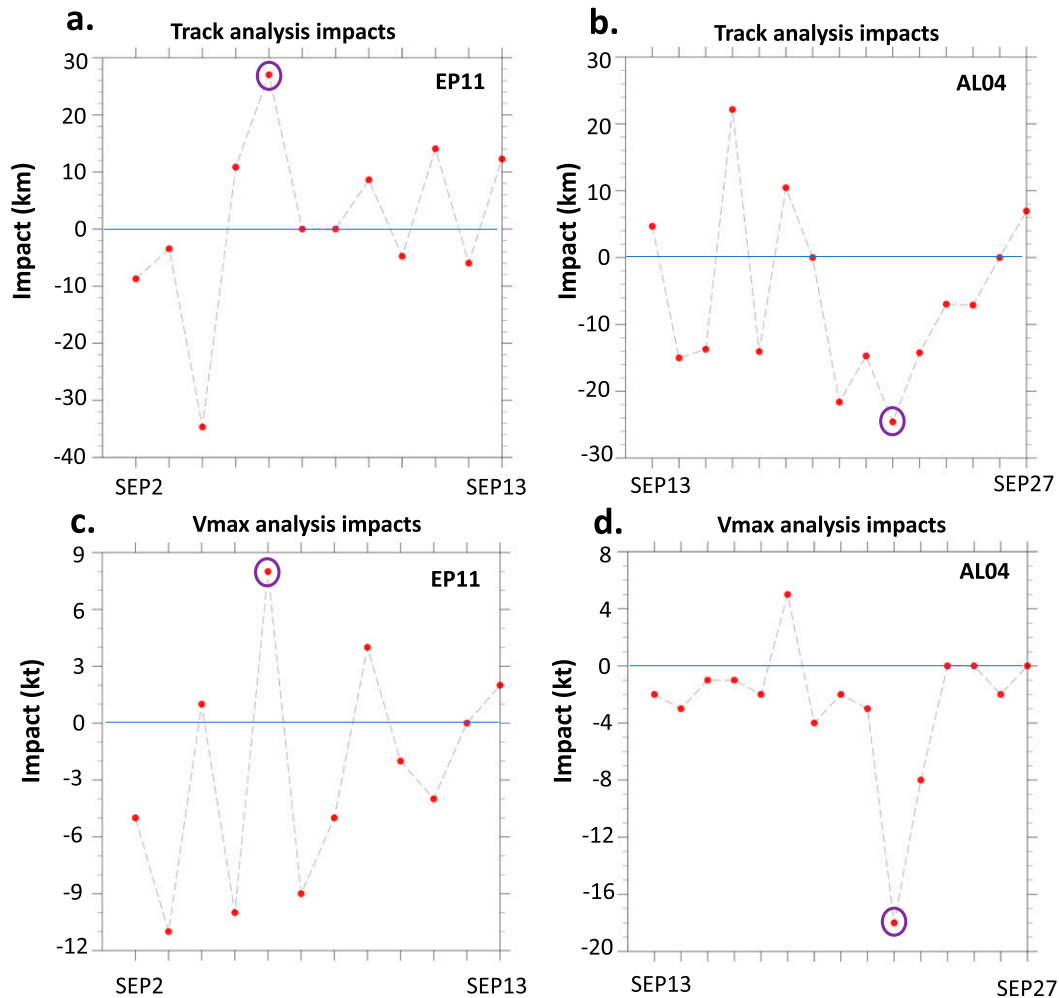


FIG. 13. Time series of GFS 0 h track and  $V_{\max}$  impact for (a),(c) EP11 and (b),(d) AL04. Impact is defined as in Fig. 3, as CTL error magnitude minus RO<sub>err</sub> error magnitude. Positive values denote RO<sub>err</sub> is improved compared to CTL. Purple circles denote the improved analysis at 0000 UTC 6 Sep for EP11 and the degraded analysis at 0000 UTC 22 Sep for AL04.

impact of GNSS-RO refractivity assimilation on TC forecasts, 2) assess the impact of GNSS-RO assimilation on global model TC forecasts to the extent that the OSSE allows, and 3) provide suggestions to improve future similar studies where necessary.

The OSSE system included a high-resolution, non-hydrostatic nature run (G5NR), an operational analysis and forecast model system (GDAS/GFS), a hybrid ensemble data assimilation system (3DVar GDAS), RO observations based on the original planned COSMIC-2 constellation, and simulated observation errors. The study was based on two 6-week experiments: One without RO assimilation (CTL) and one with RO assimilation with realistic random errors added to the RO profiles (RO<sub>err</sub>). A third experiment (RO<sub>perf</sub>) assimilated RO profiles without added errors. Supplemental one-cycle

forecasts were conducted to determine 1) the impact of a specific set of RO profiles on assimilation minimization statistics and analyses (CTL<sub>1cyc\_ro</sub>), and 2) the extent of impact originating from the background field and from the RO profiles in the current assimilation window (RO<sub>err\_1cyc\_ctl</sub>).

In general, analysis and forecast impacts are neutral to slightly negative. The major findings of the study are:

- 1) The RO refractivity assimilation produced consistent reductions in the refractivity bias and spread globally, regionally around TCs, and for the individual profiles closest to TCs. Assimilation statistics near TCs appeared to be very similar to those away from TCs and globally. Thus, there is high confidence RO is assimilated well (though not necessarily optimally) in the vicinity of TCs.

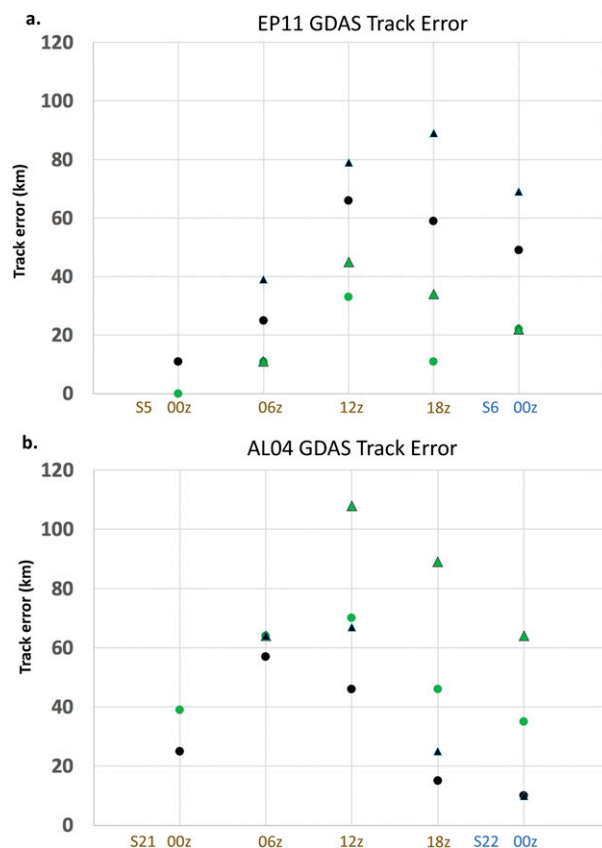


FIG. 14. Time series of track error in GDAS background (triangle) and analysis (circle) states for the four cycles preceding (a) the 0000 UTC 6 Sep EP11 analysis and (b) the 0000 UTC 22 Sep AL04 analysis. Black markers denote CTL error and green markers denote RO\_err error. Lower values represent better analyses and forecasts. The background states denoted as triangles are 6-h forecasts initialized from the previous cycle's analysis.

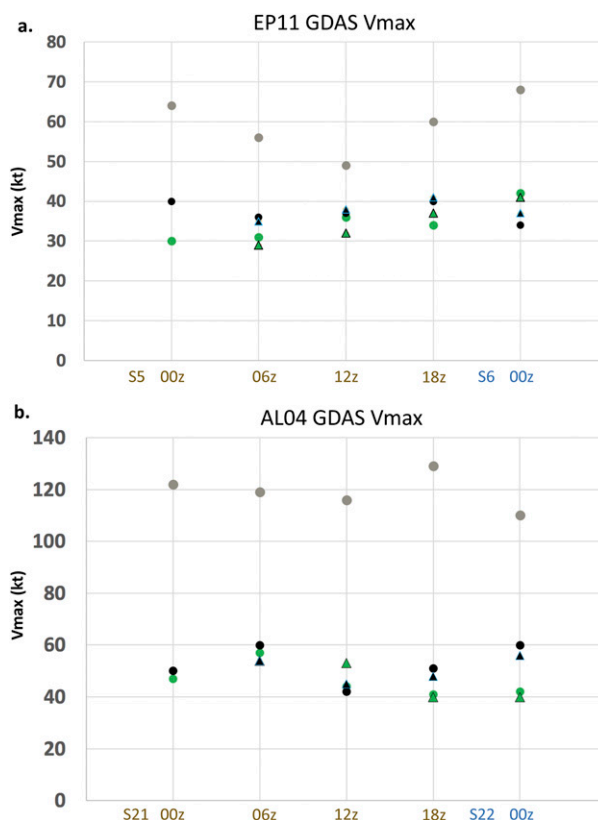


FIG. 15. Time series of GDAS  $V_{\max}$  background (triangle) and analysis (circle) states for the four cycles preceding (a) the 0000 UTC 6 Sep EP11 analysis and (b) the 0000 UTC 22 Sep AL04 analysis. Black markers denote CTL  $V_{\max}$ , green markers denote RO\_err  $V_{\max}$ , and tan markers denote G5NR  $V_{\max}$ . In this case, higher values of  $V_{\max}$  denote better analyses and forecasts. The background states denoted as triangles are 6-h forecasts initialized from the previous cycle's analysis.

- 2) Adding simulated refractivity errors in the global OSSE reduced the accuracy of both TC track and  $V_{\max}$  forecasts (as expected and desired), but impacts generally did not exceed the 95% confidence interval. Thus, in this study, added errors were not critical for interpretation of TC impacts.
- 3) Overall, TC track forecasts are not significantly impacted over the 168 h forecast; however, there is large variability on a forecast-by-forecast basis (i.e., some track forecasts are improved, while others are degraded). There is high confidence that this OSSE is capable of adequately assessing track impact because GFS track errors closely matched those documented at the National Hurricane Center.
- 4) Overall, TC  $V_{\max}$  forecasts are degraded by  $\sim 1\text{--}2$  kt over the first 60 h of the forecast. Degradations do not increase over time, suggesting RO is not degrading the near-TC environment in which the storms develop. Rather, the degradations are introduced in the

- analysis of the storm itself and begin to fade after 48 h. Confidence in the metric itself is low due to large CTL and RO\_err errors compared to official NHC errors (i.e.,  $V_{\max}$  errors are too large in part due to coarse model resolution). The value of this result is that it suggests small systematic degradation in the structure of the storms. Further analysis confirmed corresponding impact in TC structure.
- 5) Integrated kinetic energy (IKE) forecasts were slightly degraded in the first 30 h of lead time and impacts were neutral thereafter. The initial degradation did not grow with time, similar to  $V_{\max}$ . The similarity of IKE and  $V_{\max}$  results increases our confidence in concluding early small degradations for storm intensity due to refractivity assimilation.
- 6) Insensitivity of track and  $V_{\max}$  forecasts to TC geographic location, TC intensity at initialization, or the number of assimilated RO observations in the near-storm environment.

- 7) For the two largest  $V_{\max}$  analysis impacts, most of the impact originates in the background field, not the observations in the current assimilation cycle. While this is not surprising, it complicates the search for a definitive cause of RO impacts. In such cases, it is valuable to evaluate the series of preceding cycles to identify how the impact originated. Closer evaluation of analysis cycles immediately preceding two cases reveals the importance of the propagation of assimilation impact in the 6 h GDAS forecast.

It is important to consider the real-world context in which this study was conducted: Official forecasts and TC best track databases report track in  $0.1^\circ$  increments and  $V_{\max}$  in 5 kt increments. The average statistically significant impacts of RO found in this study are smaller than these increments. Thus, RO from COSMIC-2 (used this way in a global model) is not likely to have a consistent impact on forecasts issued by the NHC and consumed by the public.

As discussed in section 1, Huang et al. (2010) and Chen et al. (2015) discovered largely neutral impacts on track and did not evaluate any TC intensity metric. Chen et al. (2015) found a statistically significant improvement of 12 km at 72 h. This study generally agrees with these neutral results. The lack of improvement in track forecasts is not unique to RO assimilation studies—Landsea and Cangialosi (2018) documented stagnation in track forecast improvements during the 2010s. As they mention, if a deterministic predictability limit has been reached, future improvements will likely come from probabilistic forecasts.

Forecast improvements for  $V_{\max}$  have been even more difficult to achieve. Previous studies that have documented RO impact on TC intensity metrics have used case studies in regional limited area models and are difficult to compare to this study.  $V_{\max}$  is also a difficult metric to interpret in a global OSSE because of the coarse model involved. In a real TC,  $V_{\max}$  occurs in a very small part of the storm (i.e., the eyewall for more developed storms), driven by processes not resolved in the model used here. Future data impact studies using global models could evaluate intensity through 1) using the operational resolution of the global modeling system, 2) spatially integrated metrics like IKE, or 3) by using the global model output as boundary conditions for higher-resolution regional models designed for convection and TCs. The first option may not be feasible depending on computing resources and OSSE studies should use caution with the third option since the experimental forecast model resolution should never exceed that of the nature run. For that reason, special OSSE configurations have been developed specifically for TCs, based on nature

runs with much higher resolution (e.g., Nolan et al. 2013; Atlas et al. 2015).

It is important to note that the way observations impact the analyses and forecasts is dependent upon the data assimilation method used. In this study, we used the 3DVar GDAS. Further study is needed to determine how different methods may handle GNSS-RO data differently. Further OSSEs adding an economic component to TC impact are also recommended because TCs often cause extensive economic damage.

Furthermore, in this study the data assimilation system closely parallels the system that was used operationally a few years ago. That does not guarantee that the data assimilation system was correctly tuned to maximize the benefit of RO observations for TC forecasting. While the analysis of the  $O - B$  and  $O - A$  statistics (Figs. 7 and 8) shows the data assimilation system is better fitting the RO observations, it does not prove the data assimilation is optimal for these observations. For example, the localization scales were tuned by NOAA's Environmental Modeling Center, not specially designed for our configuration. This may have degraded the analysis of other fields. These issues should be the subject of future studies.

To best confirm this study's conclusions, a cycled global model OSE with real-world COSMIC-2 data will need to be conducted. This will soon be possible because the equatorial component of the constellation (COSMIC-2A) launched in June 2019. It is also necessary to evolve OSSE systems with the most recent modeling systems and applications. To that end, a new FV3-GFS-based OSSE system is being developed.

**Acknowledgments.** The authors thank Ruifang Li for code to plot RO profile locations, Jack Woollen for providing templates to simulate COSMIC-2 observations, Lisa Bucci for access to her integrated kinetic energy code, and to the rest of the NOAA OSSE team for their help simulating observations from the nature run. We obtained the real-world observation system templates for August–September 2014 from the NOAA/Environmental Modeling Center. The authors also thank Altug Aksoy and four anonymous reviewers for providing helpful comments to greatly improve the manuscript. This study was conducted by NOAA in response to the H.R. 353, the “Weather Research and Forecasting Innovation Act of 2017” (Public Law 115-25, <https://www.congress.gov/bill/115th-congress/house-bill/353/text>). According to section 107(d)(1) of the Act: “Not later than 30 days after the date of the enactment of this Act, the Assistant Administrator for Oceanic and Atmospheric Research shall complete an Observing System Simulation Experiment [OSSE] to

assess the value of data from Global Navigation Satellite System Radio Occultation.” Michael J. Mueller was supported by funding from NOAA Award Number NA17OAR4320101.

## REFERENCES

- Aberson, S. D., and M. DeMaria, 1994: Verification of a nested barotropic hurricane track forecast model (VICBAR). *Mon. Wea. Rev.*, **122**, 2804–2815, [https://doi.org/10.1175/1520-0493\(1994\)122<2804:VOANBH>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<2804:VOANBH>2.0.CO;2).
- Anisetty, S. K. A. V. P. R., C.-Y. Huang, and S.-Y. Chen, 2014: Impact of FORMOSAT-3/COSMIC radio occultation data on the prediction of super cyclone Gonu (2007): A case study. *Nat. Hazards*, **70**, 1209–1230, <https://doi.org/10.1007/s11069-013-0870-0>.
- Anlauf, H., D. Pingel, and A. Rhodin, 2011: Assimilation of GPS radio occultation data at DWD. *Atmos. Meas. Tech.*, **4**, 1105–1113, <https://doi.org/10.5194/amt-4-1105-2011>.
- Anthes, R. A., and Coauthors, 2008: The COSMIC/FORMOSAT-3 mission: Early results. *Bull. Amer. Meteor. Soc.*, **89**, 313–334, <https://doi.org/10.1175/BAMS-89-3-313>.
- Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan*, **75**, 111–130, [https://doi.org/10.2151/jmsj1965.75.1B\\_111](https://doi.org/10.2151/jmsj1965.75.1B_111).
- , L. Bucci, B. Annane, R. Hoffman, and S. Murillo, 2015: Observing system simulation experiments to assess the potential impact of new observing systems on hurricane forecasting. *Mar. Technol. Soc. J.*, **49**, 140–148, <https://doi.org/10.4031/MTSJ.49.6.3>.
- Bauer, P., G. Radnóti, S. Healy, and C. Cardinali, 2014: GNSS radio occultation constellation observing system experiments. *Mon. Wea. Rev.*, **142**, 555–572, <https://doi.org/10.1175/MWR-D-13-00130.1>.
- Biondi, R., A. K. Steiner, G. Kirchengast, and T. Rieckh, 2015: Characterization of thermal structure and conditions for overshooting of tropical and extratropical cyclones with GPS radio occultation. *Atmos. Chem. Phys.*, **15**, 5181–5193, <https://doi.org/10.5194/acp-15-5181-2015>.
- Boukabara, S.-A., and Coauthors, 2016: Community Global Observing System Simulation Experiment (OSSE) package (CGOP): Description and usage. *J. Atmos. Oceanic Technol.*, **33**, 1759–1777, <https://doi.org/10.1175/JTECH-D-16-0012.1>.
- Buontempo, C., A. Jupp, and M. Rennie, 2008: Operational NWP assimilation of GPS radio occultation data. *Atmos. Sci. Lett.*, **9**, 129–133, <https://doi.org/10.1002/asl.173>.
- Businger, S., and Coauthors, 1996: The promise of GPS in atmospheric monitoring. *Bull. Amer. Meteor. Soc.*, **77**, 5–18, [https://doi.org/10.1175/1520-0477\(1996\)077<0005:TPOGIA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0005:TPOGIA>2.0.CO;2).
- Chen, Y.-C., M.-E. Hsieh, L.-F. Hsiao, Y.-H. Kuo, M.-J. Yang, C.-Y. Huang, and C.-S. Lee, 2015: Systematic evaluation of the impacts of GPSRO data on the prediction of typhoons over the Northwestern Pacific in 2008–2010. *Atmos. Meas. Tech.*, **8**, 2531–2542, <https://doi.org/10.5194/amt-8-2531-2015>.
- Cucurull, L., 2010: Improvement in the use of an operational constellation of GPS radio occultation receivers in weather forecasting. *Wea. Forecasting*, **25**, 749–767, <https://doi.org/10.1175/2009WAF2222302.1>.
- , and R. A. Anthes, 2014: Impact of infrared, microwave, and radio occultation satellite observations on operational numerical weather prediction. *Mon. Wea. Rev.*, **142**, 4164–4186, <https://doi.org/10.1175/MWR-D-14-00101.1>.
- , J. C. Derber, R. Treadon, and R. J. Purser, 2007: Assimilation of global positioning system radio occultation observations into NCEP’s global data assimilation system. *Mon. Wea. Rev.*, **135**, 3174–3193, <https://doi.org/10.1175/MWR3461.1>.
- , R. Li, and T. R. Peevey, 2017: Assessment of radio occultation observations from the COSMIC-2 mission with a simplified observing system simulation experiment configuration. *Mon. Wea. Rev.*, **145**, 3581–3597, <https://doi.org/10.1175/MWR-D-16-0475.1>.
- , R. Atlas, R. Li, M. J. Mueller, and R. N. Hoffman, 2018: An observing system simulation experiment with a constellation of radio occultation satellites. *Mon. Wea. Rev.*, **146**, 4247–4259, <https://doi.org/10.1175/MWR-D-18-0089.1>.
- Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.*, **116**, 140–158, <https://doi.org/10.1016/j.rse.2010.10.017>.
- Errico, R. M., R. Yang, N. C. Privé, K.-S. Tai, R. Todling, M. E. Sienkiewicz, and J. Guo, 2013: Development and validation of observing-system simulation experiments at NASA’s Global Modeling and Assimilation Office. *Quart. J. Roy. Meteor. Soc.*, **139**, 1162–1178, <https://doi.org/10.1002/qj.2027>.
- Gelaro, R., and Coauthors, 2014: Evaluation of the 7-km GEOS-5 nature run. NASA Tech. Rep. Series on Global Modeling and Data Assimilation, TM-2014-104606, NASA Goddard Space Flight Center, 305 pp., [gmao.gsfc.nasa.gov/pubs/docs/Gelaro736.pdf](https://gmao.gsfc.nasa.gov/pubs/docs/Gelaro736.pdf).
- Gilleland, E., A. S. Hering, T. L. Fowler, and B. G. Brown, 2018: Testing the tests: What are the impacts of incorrect assumptions when applying confidence intervals or hypothesis tests to compare competing forecasts? *Mon. Wea. Rev.*, **146**, 1685–1703, <https://doi.org/10.1175/MWR-D-17-0295.1>.
- Hajj, G. A., E. R. Kursinski, L. J. Romans, W. I. Bertiger, and S. S. Leroy, 2002: A technical description of atmospheric sounding by GPS occultation. *J. Atmos. Sol.-Terr. Phys.*, **64**, 451–469, [https://doi.org/10.1016/S1364-6826\(01\)00114-6](https://doi.org/10.1016/S1364-6826(01)00114-6).
- Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155–167, [https://doi.org/10.1175/1520-0434\(1999\)014<0155:HTFENP>2.0.CO;2](https://doi.org/10.1175/1520-0434(1999)014<0155:HTFENP>2.0.CO;2).
- , J. S. Whitaker, M. Fiorino, and S. G. Benjamin, 2011: Global ensemble predictions of 2009’s tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.*, **139**, 668–688, <https://doi.org/10.1175/2010MWR3456.1>.
- Healy, S. B., 2008: Forecast impact experiment with a constellation of GPS radio occultation receivers. *Atmos. Sci. Lett.*, **9**, 111–118, <https://doi.org/10.1002/asl.169>.
- Ho, S., X. Yue, Z. Zeng, C. O. Ao, C. Huang, E. R. Kursinski, and Y. Kuo, 2014: Applications of COSMIC radio occultation data from the troposphere to ionosphere and potential impacts of COSMIC-2 data. *Bull. Amer. Meteor. Soc.*, **95**, ES18–ES22, <https://doi.org/10.1175/BAMS-D-13-00035.1>.
- Hoffman, R. N., and R. Atlas, 2016: Future observing system simulation experiments. *Bull. Amer. Meteor. Soc.*, **97**, 1601–1616, <https://doi.org/10.1175/BAMS-D-15-00200.1>.
- Huang, C.-Y., Y. Kuo, S. Chen, and F. Vandenbergh, 2005: Improvements in typhoon forecasts with assimilated GPS occultation refractivity. *Wea. Forecasting*, **20**, 931–953, <https://doi.org/10.1175/WAF874.1>.
- , and Coauthors, 2010: Impact of GPS radio occultation data assimilation on regional weather predictions. *GPS Solutions*, **14**, 35–49, <https://doi.org/10.1007/s10291-009-0144-1>.
- Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation, and Predictability*. Cambridge University Press, 341 pp., <https://doi.org/10.1017/CBO9780511802270>.



- Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part I: System description and 3D-Hybrid results. *Mon. Wea. Rev.*, **143**, 433–451, <https://doi.org/10.1175/MWR-D-13-00351.1>.
- Kren, A. C., L. Cucurull, and H. Wang, 2018: Impact of UAS global hawk dropsonde data on tropical and extratropical cyclone forecasts in 2016. *Wea. Forecasting*, **33**, 1121–1141, <https://doi.org/10.1175/WAF-D-18-0029.1>.
- Kuo, Y.-H., H. Liu, Y.-R. Guo, C.-T. Terng, and Y.-T. Lin, 2009: Impact of FORMOSAT-3/COSMIC data on typhoon and Mei-yu prediction. *Recent Progress in Atmospheric Sciences: Applications to the Asia-Pacific Region*, World Scientific, 458–483.
- Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, **102**, 23 429–23 465, <https://doi.org/10.1029/97JD01569>.
- Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, <https://doi.org/10.1175/MWR-D-12-00254.1>.
- , and J. P. Cangialosi, 2018: Have we reached the limits of predictability for tropical cyclone track forecasting? *Bull. Amer. Meteor. Soc.*, **99**, 2237–2243, <https://doi.org/10.1175/BAMS-D-17-0136.1>.
- Leidner, S. M., T. Nehrkorn, J. Henderson, M. Mountain, T. Yunck, and R. N. Hoffman, 2017: A severe weather Quick Observing System Simulation Experiment (QuickOSSE) of Global Navigation Satellite System (GNSS) Radio Occultation (RO) superconstellations. *Mon. Wea. Rev.*, **145**, 637–651, <https://doi.org/10.1175/MWR-D-16-0212.1>.
- Li, X., F. Zus, C. Lu, T. Ning, G. Dick, M. Ge, J. Wickert, and H. Schuh, 2015: Retrieving high-resolution tropospheric gradients from multiconstellation GNSS observations. *Geophys. Res. Lett.*, **42**, 4173–4181, <https://doi.org/10.1002/2015GL063856>.
- Liu, H., J. Anderson, and Y. Kuo, 2012: Improved analyses and forecasts of Hurricane Ernesto's genesis using radio occultation data in an ensemble filter assimilation system. *Mon. Wea. Rev.*, **140**, 151–166, <https://doi.org/10.1175/MWR-D-11-00024.1>.
- Murphy, B. J., J. S. Haase, P. Muradyan, J. L. Garrison, and K.-N. Wang, 2015: Airborne GPS radio occultation refractivity profiles observed in tropical storm environments. *J. Geophys. Res. Atmos.*, **120**, 1690–1709, <https://doi.org/10.1002/2014JD022931>.
- NOAA, 2019a: Average NHC Atlantic intensity forecast errors. NOAA, 2 pp., [https://www.nhc.noaa.gov/verification/pdfs/1990-present\\_OFCL\\_ATL\\_annual\\_int\\_errors.pdf](https://www.nhc.noaa.gov/verification/pdfs/1990-present_OFCL_ATL_annual_int_errors.pdf).
- , 2019b: Average NHC Atlantic track forecast errors. NOAA, 2 pp., [https://www.nhc.noaa.gov/verification/pdfs/1989-present\\_OFCL\\_ATL\\_annual\\_trk\\_errors.pdf](https://www.nhc.noaa.gov/verification/pdfs/1989-present_OFCL_ATL_annual_trk_errors.pdf).
- Nolan, D. S., R. M. Atlas, K. T. Bhatia, and L. R. Bucci, 2013: Development and validation of a hurricane nature run using the joint OSSE nature run and WRF model. *J. Adv. Model. Earth Syst.*, **5**, 382–405, <https://doi.org/10.1002/jame.20031>.
- NWS, 2014: Corrected: Global Forecast Systems (GFS) update: Effective January 14, 2015. Tech. Implementation Notice TIN14-46, National Weather Service, accessed 13 October 2019, [https://www.weather.gov/media/notification/tins/tin14-46gfs\\_aab.pdf](https://www.weather.gov/media/notification/tins/tin14-46gfs_aab.pdf).
- Phunthirawurthi, P., F. Wu, and P. Boonyuen, 2016: Performance analysis of GPS radio occultation assimilation for tropical cyclone monitoring. *Geoscience and Remote Sensing Symp. (IGARSS) 2016*, Beijing, China, IEEE, <https://doi.org/10.1109/IGARSS.2016.7729563>.
- Powell, M. D., and T. A. Reinhold, 2007: Tropical cyclone destructive potential by integrated kinetic energy. *Bull. Amer. Meteor. Soc.*, **88**, 513–526, <https://doi.org/10.1175/BAMS-88-4-513>.
- Putman, W. M., A. Darmanov, A. da Silva, R. Gelaro, A. Molod, L. Ott, and M. J. Suarez, 2015: A 7-km non-hydrostatic global mesoscale simulation for OSSEs with the Goddard Earth Observing System model (GEOS-5). *19th Conf. on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)*, Phoenix, AZ, Amer. Meteor. Soc., 3.1, <https://ams.confex.com/ams/95Annual/webprogram/Paper260701.html>.
- Reale, O., D. Achuthavarier, M. Fuentes, W. M. Putman, and G. Partyka, 2017: Tropical cyclones in the 7-km NASA global nature run for use in observing system simulation experiments. *J. Atmos. Oceanic Technol.*, **34**, 73–100, <https://doi.org/10.1175/JTECH-D-16-0094.1>.
- Vergados, P., Z. J. Luo, K. Emanuel, and A. J. Mannucci, 2014: Observational tests of hurricane intensity estimations using GPS radio occultations. *J. Geophys. Res. Atmos.*, **119**, 1936–1948, <https://doi.org/10.1002/2013JD020934>.
- Wang, X., 2010: Incorporating ensemble covariance in the grid-point statistical interpolation variational minimization: A mathematical framework. *Mon. Wea. Rev.*, **138**, 2990–2995, <https://doi.org/10.1175/2010MWR3245.1>.
- , D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble-variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, <https://doi.org/10.1175/MWR-D-12-00141.1>.
- Weather Research and Forecasting Innovation Act, 2017: Public Law 115-25. 131 STAT. 91, 38 pp., <https://www.congress.gov/115/plaws/publ25/PLAW-115publ25.pdf>.